CURRENT PROTOCOLS IN PROTEIN SCIENCE

VOLUME 1

EDITED BY

JOHN E. COLIGAN
BEN M. DUNN
HIDDE L. PLOEGH
DAVID W. SPEICHER
PAUL T. WINGFIELD

Masses and Compositions of the Twenty Commonly Occurring Amino Acid Residues

Name	Abbreviations		Residue Structure	Average Mass[a]	pK_a of Ionizing Side Chain[b]	Occurrence in Proteins (%)
Alanine	Ala	A		71.0788	—	8.3
Arginine	Arg	R		156.1876	~12	5.7
Asparagine	Asn	N		114.1039	—	4.4
Aspartic Acid	Asp	D		115.0886	3.9–4.0	5.3
Cysteine	Cys	C		103.1448	9.0–9.5	1.7
Glutamic Acid	Glu	E		129.1155	4.3–4.5	6.2
Glutamine	Gln	Q		128.1308	—	4.0
Glycine	Gly	G		57.0520	—	7.2
Histidine	His	H		137.1412	6.0–7.0	2.2
Isoleucine	Ile	I		113.1595	—	5.2
Leucine	Leu	L		113.1595	—	9.0
Lysine	Lys	K		128.1742	10.4–11.1	5.7
Methionine	Met	M		131.1986	—	2.4
Phenylalanine	Phe	F		147.1766	—	3.9
Proline	Pro	P		97.1167	—	5.1
Serine	Ser	S		87.0782	—	6.9
Threonine	Thr	T		101.1051	—	5.8
Tryptophan	Trp	W		186.2133	—	1.3
Tyrosine	Tyr	Y		163.1760	10.0–10.3	3.2
Valine	Val	V		99.1326	—	6.6

[a] The molecular mass of a normally terminated and unmodified peptide or protein may be calculated by summing the masses of the appropriate amino acid residues and adding the masses of H and OH for the N and C termini, respectively. In cases where cysteines are linked by disulfide bridges, the mass of 2 hydrogen atoms should be subtracted for each disulfide bridge in the molecule.

[b] These values are included for anyone wishing to make approximate isoelectric point determinations based on protein composition. Values for the terminal residues depend on the identity of the residue: α-amino, pK_a 6.8–8.0; α-carboxyl, pK_a 3.5–4.3.

22.843 CN-1946

CURRENT PROTOCOLS IN PROTEIN SCIENCE

VOLUME 1

X,[1010]P

BCU IASI / CENTRAL UNIVERSITY LIBRARY

EDITORIAL BOARD

John E. Coligan
National Institute of Allergy and Infectious Diseases
Bethesda, Maryland

Ben M. Dunn
University of Florida
Gainesville, Florida

Hidde L. Ploegh
Harvard Medical School
Boston, Massachusetts

David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

Paul T. Wingfield
National Institute of Arthritis and Musculoskeletal and Skin Diseases
Bethesda, Maryland

SERIES EDITOR

Virginia Benson Chanda

054758
B.C.U. - IASI

John Wiley & Sons, Inc.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

CORE 4 (S12)

BIBL. CENTR. UNIV.
„M. EMINESCU" IAŞI
C. I. M. I.
C IV 1946

Copyright © 1995–1998 by John Wiley & Sons, Inc.

All rights reserved.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

While the authors, editors, and publisher believe that the specification and usage of reagents, equipment, and devices, as set forth in this book, are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. This is particularly important in regard to new or infrequently employed chemicals or experimental reagents.

Library of Congress Cataloging in Publication Data:

Current protocols in protein science / editorial board, John E. Coligan ... [et al.].
p. cm.
Kept up to date by quarterly supplements.
Includes index.
ISBN 0-471-11184-8 (v. 1)
1. Proteins—Laboratory manuals. I. Coligan, John E.

QP551.C86 1995
574.19'245—dc20

95-1030
CIP

Printed in the United States of America

10 9 8 7 6 5 4

22843
B.C.U. „M. EMINESCU" IAŞI
17 III 2000

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Current Protocols Order Form
15% Subscriber Discount

For your convenience, you may photocopy this form and fax your order:

★ 1-212-850-6021, Attn: Current Protocols Customer Service

You may also place your order by:

Telephone: 1-800-825-7550 or 212-850-6347 *E-Mail:* protocol@wiley.com

Mail: John Wiley & Sons, Inc. / Attn: Current Protocols Customer Service / 605 Third Avenue, 9th Floor / New York, NY 10158 USA

Prices include 15% discount and a one-year subscription to the quarterly update service:

NOTE: Shipping and handling for one-year service must be added to all orders:
Looseleaf - $50 per order
CD-ROM - $40 per disc *(All publications are included on one disc. If you currently subscribe to CP on CD-ROM, you do not need to pay shipping for adding a publication.)*

LOOSELEAF SERVICE with 1-year quarterly updates

- ❑ **Molecular Biology** $454.75
- ❑ **Immunology** $454.75
- ❑ **Human Genetics** $382.50
- ❑ **Protein Science** $335.75
- ❑ **Cytometry** $250.75
- ❑ **Neuroscience** $293.25
- ❑ **Pharmacology** $335.75
- ❑ **Cell Biology** $293.25 *(Coming October '98)*

CD-ROM SERVICE with 1-year quarterly updates

- ❑ **Molecular Biology** $454.75
- ❑ **Immunology** $454.75
- ❑ **Human Genetics** $382.50
- ❑ **Protein Science** $382.50
- ❑ **Cytometry** $382.50
- ❑ **Neuroscience** $382.50
- ❑ **Pharmacology** $382.50
- ❑ **Cell Biology** $382.50 *(Coming October '98)*

SHIP MY ORDER TO *(please print or type)*

Account Number *(if known)* _____

Name _____

Institution _____

Address _____

_____

_____

Country _____

Telephone # _____

Fax # _____

TOTAL ORDER

Total Publications $_____

Shipping *($50 per Looseleaf $40 per CD disc)* $_____

Tax $_____

Grand Total $_____

METHOD OF PAYMENT

Please add appropriate state and local sales tax. In Canada, add GST. Checks are payable to John Wiley & Sons, Inc. in U.S. dollars drawn on a U.S. bank.

❑ Bill me—*This option is available to U.S. customers only.*

❑ Bill my enclosed P.O.# _____

❑ Check enclosed

❑ MasterCard ❑ VISA ❑ AMEX ❑ Discover

Card number _____

Expiration date _____

Signature _____

EG161

Additional Information:

In Japan: All orders must be placed with Maruzen.

30-Day Examination: All Current Protocols publications can be examined for 30 days at no charge. If for any reason you decide it does not meet your needs, you may return the product without further obligation and receive a full refund. Prepayment is required outside of the U.S.

CD-ROM Subscriptions: Each quarterly disc is programmed to expire after the date of issuance of the next update disc. At the end of the one-year license, subscribers who decide not to renew the CD-ROM Service may elect to receive an archive version of the program for a nominal fee (not to exceed one-quarter of the annual subscription renewal fee).

For more information about all Current Protocols publications, please refer to our Home Page:
www.wiley.com/cp

CURRENT CP PROTOCOLS

Current Protocols is published by John Wiley & Sons, Inc.

WILEY

BOU IAS CENTRAL UNIVERSITY LIBRARY

Table of Contents, Volumes 1 and 2

i Foreword

iii Preface

vii Contributors

1 Strategies of Protein Purification and Characterization

1.1 Overview of Protein Purification and Characterization
1.2 Strategies for Protein Purification
1.3 Protein Purification Flow Charts

2 Computational Analysis

2.1 Computational Methods for Protein Sequence Analysis
2.2 Hydrophobicity Profiles for Protein Sequence Analysis
2.3 Protein Secondary Structure Prediction
2.4 Internet Basics
2.5 Sequence Similarity Searching Using the BLAST Family of Programs

3 Detection and Assay Methods

3.1 Spectrophotometric Determination of Protein Concentration
3.2 Quantitative Amino Acid Analysis
3.3 In Vitro Radiolabeling of Peptides and Proteins
3.4 Assays for Total Protein
3.5 Kinetic Assay Methods
3.6 Biotinylation of Proteins in Solution and on Cell Surfaces

4 Extraction, Stabilization, and Concentration

4.1 Overview of Cell Fractionation
4.2 Purification of Organelles from Mammalian Cells
4.3 Subcellular Fractionation of Tissue Culture Cells
4.4 Desalting, Concentration, and Buffer Exchange by Dialysis and Ultrafiltration
4.5 Selective Precipitation of Proteins

5 Production of Recombinant Proteins

5.1 Production of Recombinant Proteins in *Escherichia coli*
5.2 Selection of *Escherichia coli* Expression Systems
5.3 Fermentation and Growth of *Escherichia coli* for Optimal Protein Production
5.4 Overview of the Baculovirus Expression System
5.5 Protein Expression in the Baculovirus System
5.6 Overview of Protein Expression in *Saccharomyces cerevisiae*

continued

Copyright © 1998 by John Wiley & Sons, Inc.

BCU MAIN CENTRAL UNIVERSITY LIBRARY

5.7 Overview of Protein Expression in *Pichia pastoris*
5.8 Culture of Yeast for the Production of Heterologous Proteins
5.9 Overview of Protein Expression by Mammalian Cells
5.10 Production of Recombinant Proteins in Mammalian Cells

6 Purification of Recombinant Proteins

6.1 Overview of the Purification of Recombinant Proteins Produced in *Escherichia coli*
6.2 Preparation of Soluble Proteins from *Escherichia coli*
6.3 Preparation and Extraction of Insoluble (Inclusion-Body) Proteins from *Escherichia coli*
6.4 Overview of Protein Folding
6.5 Folding and Purification of Insoluble (Inclusion-Body) Proteins from *Escherichia coli*
6.6 Expression and Purification of GST Fusion Proteins
6.7 Expression and Purification of Thioredoxin Fusion Proteins

7 Characterization of Recombinant Proteins

7.1 Overview of the Characterization of Recombinant Proteins
7.2 Determining the Identity and Purity of Recombinant Proteins by UV Absorption Spectroscopy
7.3 Determining the Identity and Structure of Recombinant Proteins
7.4 Transverse Urea-Gradient Gel Electrophoresis
7.5 Analytical Ultracentrifugation
7.6 Determining the CD Spectrum of a Protein
7.7 Determining the Fluorescence Spectrum of a Protein
7.8 Light Scattering
7.9 Measuring Protein Thermostability by Differential Scanning Calorimetry

8 Conventional Chromatographic Separations

8.1 Overview of Conventional Chromatography
8.2 Ion-Exchange Chromatography
8.3 Gel-Filtration Chromatography
8.4 Hydrophobic-Interaction Chromatography
8.5 Chromatofocusing
8.6 Hydroxylapatite Chromatography

9 Affinity Purification

9.1 Lectin Affinity Chromatography
9.2 Dye Affinity Chromatography
9.3 Affinity Purification of Natural Ligands
9.4 Metal-Chelate Affinity Chromatography
9.5 Immunoaffinity Chromatography
9.6 Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography
9.7 Purification of DNA-Binding Proteins Using Biotin/Streptavidin Affinity Systems

BCU IAS / CENTRAL UNIVERSITY LIBRARY

10 Electrophoresis

10.1 One-Dimensional SDS Gel Electrophoresis of Proteins
10.2 One-Dimensional Isoelectric Focusing of Proteins in Slab Gels
10.3 One-Dimensional Electrophoresis Using Nondenaturing Conditions
10.4 Two-Dimensional Gel Electrophoresis
10.5 Protein Detection in Gels Using Fixation
10.6 Protein Detection in Gels Without Fixation
10.7 Electroblotting from Polyacrylamide Gels
10.8 Detection of Proteins on Blot Membranes
10.9 Capillary Electrophoresis of Proteins and Peptides
10.10 Immunoblot Detection
10.11 Autoradiography
10.12 Overview of Digital Electrophoresis Analysis

11 Chemical Analysis

11.1 Enzymatic Digestion of Proteins in Solution
11.2 Enzymatic Digestion of Proteins on PVDF Membranes
11.3 Digestion of Proteins in Gels for Sequence Analysis
11.4 Chemical Cleavage of Proteins in Solution
11.5 Chemical Cleavage of Proteins on Membranes
11.6 Reversed-Phase Isolation of Peptides
11.7 Removal of N-Terminal Blocking Groups from Proteins
11.8 C-Terminal Sequence Analysis
11.9 Amino Acid Analysis
11.10 N-Terminal Sequence Analysis of Proteins and Peptides

12 Post-Translational Modification: Glycosylation

12.1 Overview of Glycoconjugate Analysis
12.2 Metabolic Radiolabeling of Animal Cell Glycoconjugates
12.3 Inhibition of N-Linked Glycosylation
12.4 Endoglycosidase and Glycoamidase Release of N-Linked Oligosaccharides
12.5 Detection of Glycophospholipid Anchors on Proteins

13 Post-Translational Modification: Phosphorylation and Phosphatases

13.1 Overview of Protein Phosphorylation
13.2 Labeling Cultured Cells with $^{32}P_i$ and Preparing Cell Lysates for Immunoprecipitation
13.3 Phosphoamino Acid Analysis
13.4 Detection of Phosphorylation by Immunological Techniques
13.5 Detection of Phosphorylation by Enzymatic Techniques
13.6 Production of Antibodies That Recognize Specific Tyrosine-Phosphorylated Peptides
13.7 Assays of Protein Kinases Using Exogenous Substrates
13.8 Permeabilization Strategies to Study Protein Phosphorylation

BCIT LIBRARY - BRITISH COLUMBIA INSTITUTE OF TECHNOLOGY

14 Post-Translational Modification: Specialized Applications

14.1 Analysis of Disulfide Bond Formation
14.2 Analysis of Protein Acylation
14.3 Analysis of Protein Prenylation and Carboxyl-Methylation

15 Chemical Modification of Proteins

15.1 Modification of Cysteine
15.2 Modification of Amino Acids

16 Mass Spectrometry

16.1 Overview of Peptide and Protein Analysis by Mass Spectrometry
16.2 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Analysis of Peptides

17 Structural Biology

17.1 Overview of Protein Structural and Functional Folds
17.2 Overview of Macromolecular Electron Microscopy
17.3 Principles of Macromolecular X-Ray Crystallography
17.4 Crystallization of Soluble Macromolecules

18 Preparation and Handling of Peptides

18.1 Introduction to Peptide Synthesis
18.2 Synthesis of Multiple Peptides on Plastic Pins
18.3 Synthetic Peptides for Production of Antibodies that Recognize Intact Proteins

Appendices

A.1 Useful Data
- A Characteristics of Amino Acids
- B Commonly Used Detergents
- C Conversion Factors and Half-Life Information for Radioactivity

A.2 Laboratory Guidelines, Equipment, and Stock Solutions
- A Laboratory Safety
- B Safe Use of Radioisotopes
- C Centrifuges and Rotors
- D Standard Laboratory Equipment
- E Commonly Used Reagents

A.3 Commonly Used Techniques
- A Use of Protein Folding Reagents
- B Dialysis
- C Techniques for Mammalian Cell Tissue Culture
- D Importing Biological Materials

Suppliers Appendix

Indexes

BCU IASLIC CENTRAL UNIVERSITY LIBRARY

FOREWORD

Over the past five to ten years we have witnessed a revolution in biological research which is fueled by our increasing capacity to decipher biological information of three types: the digital information embedded in DNA with its four-letter alphabet; the three-dimensional information represented by proteins, the major executors of biological function; and the four-dimensional information of complex biological systems and networks representing the temporal and spatial interaction of multiple components. The analysis of complex systems and networks in their entirety is essential for scientists to understand the molecular basis of fascinating processes such as growth, development, and differentiation, and to extract the biological meaning of emergent properties such as consciousness, memory, and the ability to learn which evolved over billions of years.

Not surprisingly, the linear nucleotide sequence in DNA has been the first type of information to be deciphered. Large-scale genetic mapping and DNA sequencing rapidly generate enormous amounts of biological information. In many ways, programs such as the human genome project represent some of the earliest attempts to understand biological complexity. Successful completion of the human genome and similar projects will pose the challenge of interpreting the information contained in billions of nucleotides and of explaining how the interplay of the products of perhaps 100,000 genes results in the myriad of biological phenotypes. Meeting these challenges will require new interdisciplinary strategies which draw from the expertise of scientists from disciplines as different as molecular biology, biochemistry, engineering, chemistry, applied mathematics, and computer science.

Protein science is central to such integrated strategies. Separation of proteins from complex mixtures, followed by structural and functional analyses, control of protein function, post-translational processing, and modification and formation of macromolecular complexes of proteins and other biomolecules are but a few examples of topics which are essential for all those scientists who attempt to interpret the linear DNA sequence in terms of the three- and four-dimensional information of complex biological systems and networks. The following chapters detail the necessary protocols for this endeavor, provided by experts in experimental protein science.

Two attractive features set this manual apart from other collections of research protocols. First, *Current Protocols in Protein Science* will be continuously updated and expanded by quarterly additions to the core edition. Secondly, this volume as well as the updates are also available on CD-ROM. Numerous cross-references within the manual by hypertext links, context based searching, and provisions for making individualized notebooks containing frequently used protocols are attractions of the CD-ROM version which are in tune with the increasing dependence on computers in biological research laboratories.

Protein science is a diverse, experimentally challenging, and rapidly evolving discipline. This book provides expert guidance in experiment design and execution and promises to remain up-to-date, in contents and in format, for many years to come.

Leroy E. Hood and Ruedi Aebersold
Seattle, Washington

Copyright © 1997 by John Wiley & Sons, Inc.

BCU IAS / CYFRA-UNIVERSITY LIBRARY

PREFACE

Proteins are like people: they are highly diverse, with very few traits shared by the entire population. Even close relatives often have very different structural (physical) characteristics and behave quite differently. This heterogeneity presents challenges for scientists interested in studying protein structure and function. The need for a readily accessible, updatable reference for protein methods has grown in tandem with the ongoing protein science renaissance driven by advances in molecular biology, genetics, structural biology, and related disciplines. Using recombinant methods, it is now practical to overexpress both normal and reengineered forms of a protein and to create chimeric proteins with unique new properties. The ability to express proteins usually found in low abundance at high levels and to manipulate their amino acid sequences provides unprecedented opportunities to study a vast array of proteins and their associated biological processes (both in vitro and in vivo). The combination of biochemical, biophysical, and high-resolution structural analyses of proteins with manipulation of the associated genes (in cultured cells or an entire mammalian organism) has already led to spectacular advances in many areas of biotechnology and molecular medicine.

Because proteins are so diverse, a wide range of techniques must often be considered and tested empirically for a particular protein, and each researcher's arsenal must include both recent innovations and techniques that have been used for many years. Moreover, methods appropriate for purifying and initially characterizing the protein(s) associated with a biological activity of interest are often not suitable for analyzing the corresponding overexpressed recombinant protein. This manual includes both specific detailed protocols as well as strategies for adapting methods to particular projects at hand. It is designed to meet the needs of scientists with little prior experience of protein isolation and characterization, including graduate students and scientists trained in other biological disciplines. At the same time, the broad range of techniques presented here should ensure that even seasoned experts will find new and useful approaches and will benefit from the manual's convenient compilation of standard information and methods that typically must be culled from many sources.

The quarterly updating service will allow ongoing expansion of the core volume to include chapters covering critical topics that were deferred due to initial space limitations and to keep pace with the "cutting edge" of developments in this fast-moving field.

HOW TO USE THIS MANUAL

Format and Organization

This publication is available in both looseleaf and CD-ROM format. For looseleaf purchasers, a binder is provided to accommodate the growth of the manual via the quarterly update service. The looseleaf format of the binder allows easy insertion of new pages, units, and chapters that are added. The index and table of contents are updated with each supplement. CD-ROM purchasers receive a completely new disc every quarter and should dispose of their outdated discs. The material covered in the two versions is identical.

Subjects in this manual are organized by chapters and sections, and protocols are contained in units. Units generally describe a method and include one or more protocols with listings of materials, steps and annotations, recipes for unique reagents and solutions, and commentaries on the "hows" and "whys" of the method; there are also "overview" units containing theoretical discussions that lay the foundation for subsequent protocols.

BIOLOGICAL RESEARCH LIBRARY

Page numbering in the looseleaf version reflects the modular arrangement by unit; for example, page 2.3.5 refers to Chapter 2 (Computational Analysis), *UNIT 2.3* (Protein Secondary Structure Prediction), page 5 of that particular unit.

Many reagents and procedures are employed repeatedly throughout the manual. Instead of duplicating this information, cross-references among units are used extensively. Cross-referencing helps to ensure that lengthy and complex protocols are not overburdened with steps describing auxiliary procedures needed to prepare raw materials and analyze results.

Certain units that describe commonly used techniques and recipes (e.g., gel electrophoresis, ion-exchange chromatography) are cross-referenced in other units that describe their application. Thus, whenever it is necessary to isolate or identify a protein band in a protocol, the appropriate unit in Chapter 10—describing various procedures for gel electrophoresis—is cross-referenced (i.e., *UNIT 10.1*). For some widely used techniques (such as dialysis), readers are cross-referenced to *APPENDIX 3*.

Introductory and Explanatory Information

Because this publication is first and foremost a compilation of laboratory techniques in protein science, we have not provided extensive instructive material. We have, however, included explanatory information where required to help readers gain an intuitive grasp of the procedures. Chapter 1 provides an overview of strategies for protein purification and characterization, to help the reader plan for all aspects of an experimental study. Some subsequent chapters begin with special overview units that describe the state of the art of the topic matter and provide a context for the procedures that follow. Chapter and unit introductions describe how the protocols that follow connect to one another, and annotations to the actual protocol steps describe what is happening as a procedure is carried out. Finally, the Commentary that closes each protocol unit describes background information regarding the historical and theoretical development of the method, as well as alternative approaches, critical parameters, troubleshooting guidelines, anticipated results, and time considerations. All units contain cited references and many indicate key references to inform users of particularly useful background reading, original descriptions, or applications of a technique.

Protocols

Many units in the manual contain groups of protocols, each presented with a series of steps. The *basic* protocol is presented first in each unit and is generally the recommended or most univerally applicable approach. *Alternate* protocols are provided where different equipment or reagents can be employed to achieve similar ends, where the starting material requires a variation in approach, or where requirements for the end product differ from those in the basic protocol. *Support* protocols describe additional steps that are required to perform the basic or alternate protocols; these steps are separated from the core protocol because they might be applicable to other uses in the manual, or because they are performed in a time frame separate from the basic protocol steps.

Reagents and Solutions

Reagents required for a protocol are itemized in the materials list before the procedure begins. Many are common stock solutions, others are commonly used buffers or media, while others are solutions unique to a particular protocol. Recipes for the latter solutions are provided in each unit, following the protocols (and before the commentary) under the heading Reagents and Solutions. It is important to note that the *names* of some of these special solutions might be similar from unit to unit (e.g., SDS sample buffer) while the

BCU IAS CENTRAL UNIVERSITY LIBRARY

recipes differ; thus, make certain that reagents are prepared from the proper recipes. On the other hand, recipes for commonly used stock solutions and buffers are provided once in *APPENDIX 2E*. These universal recipes are cross-referenced parenthetically in the materials lists rather than repeated with every usage.

Commercial Suppliers

Throughout the manual, we have recommended commercial suppliers of chemicals, biological materials, and equipment. In some cases, the noted brand has been found to be of superior quality or it is the only suitable product available in the marketplace. In other cases, the experience of the author of that protocol is limited to that brand. In the latter situation, recommendations are offered as an aid to the novice protein science experimenter in obtaining the tools of the trade. Experienced investigators are therefore encouraged to experiment with substituting their own favorite brands.

Addresses, phone numbers, and facsimile numbers of all suppliers mentioned in this manual are provided in the *SUPPLIERS APPENDIX*.

Safety Considerations

Anyone carrying out these protocols may encounter the following hazardous or potentially hazardous materials: (1) radioactive substances, (2) toxic chemicals and carcinogenic or teratogenic reagents, and (3) pathogenic and infectious biological agents. Readers are referred to *APPENDIX 2A* and *APPENDIX 2B* for guidelines regarding use, storage, and disposal of these materials. Although cautionary statements are included in the appropriate units with cross-references to this appendix, we emphasize that users must proceed with the prudence and precaution associated with good laboratory practice, and that all materials must be used in strict accordance with local and national regulations.

Reader Response

Most of the protocols included in this manual are used routinely in our own laboratories. These protocols work for us; to make them work for you we have annotated critical steps and included critical parameters and troubleshooting guides in the commentaries to most units. However, the successful evolution of this manual depends upon readers' observations and suggestions. Consequently, a self-mailing reader-response survey can be found at the back of the manual (and is included with each supplement); we encourage readers to send in their comments.

ACKNOWLEDGMENTS

This manual is the product of dedicated efforts by many of our scientific colleagues who are acknowledged in each unit and by the hard work by the Current Protocols editorial staff at John Wiley and Sons. The publisher's commitment and continuing support for a protein science manual, particularly through the efforts of Brian Crawford and Paul Majchrzyk, was essential for realizing this ambitious project. We are extremely grateful for the critical contributions by Virginia Chanda (Series Editor) who kept the editors and the contributors on track and played a key role in bringing the entire project to completion. Other skilled members of the Current Protocols staff who contributed to the project include Janet Blair, Hazel Chan, Elizabeth Konkle, Alice Ro, and Kathy Wisch. The extensive copyediting required to produce an accurate protocols manual was ably handled by Rebecca Barr, Karen Hopkin, Arline Keithe, Kathy Morgan, Connie Parks, and Allen Ranz, and electronic illustrations were prepared by Gae Xavier. Finally, we would like to thank two people who played important roles in initiating this project. Sarah Greene

BCU INSTRUMENTAL UNIVERSITY LIBRARY

initially conceived this project and had a critical role in developing it. John A. Smith made important contributions during the early planning phases of the project.

Helpful advice and criticisms of the project during early planning stages were provided by Fred Ausubel, Charles S. Craik, Bruce A. Cummingham, Mark Hermodson, Tony E. Hugli, Robert Kingston, Albert Light, Paul Matsudaira, David Moore, George Rose, and Kevin Struhl. Mark Hermodson read and critiqued the entire manual and individual chapters were thoughtfully reviewed by John Freed and John Kay. Finally, we are especially grateful to our many colleagues, who contributed material to this manual, for sharing their procedures and experiences.

RECOMMENDED BACKGROUND READING

Branden, C. and Tooze, J. 1991. Introduction to Protein Structure. Garland Publishing, New York.

Easy-to-read overview of basic structural principles of proteins with extensive illustrations.

Jagow, G. and Schagger, H. 1994. A Practical Guide to Membrane Protein Purification. Academic Press, San Diego.

Deals with specific problems encountered when purifying membrane-associated proteins.

Kyte, J. 1994. Structure in Protein Chemistry. Garland Publishing, New York.

Comprehensive overview of structural principles of proteins with emphasis on physical chemistry approaches.

Scopes, R.K. 1994. Protein Purification: Principles and Practice. Springer-Verlag, New York.

In-depth description of most conventional protein purification methods.

Zubay, G.L., Parson, W.W., and Vance. D.E. 1995. Principles of Biochemistry. William C. Brown Publishers, Dubuque, Iowa.

A current biochemistry textbook whose protein-related chapters provide a comprehensive background for readers with no formal training in biochemistry.

John E. Coligan, Ben M. Dunn,
Hidde L. Ploegh, David W. Speicher,
and Paul T. Wingfield

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CONTRIBUTORS

R. Donald Allison
University of Florida
Gainesville, Florida

Fernando Aniento
University of Geneva Sciences II
Geneva, Switzerland

Ronald S. Annan
SmithKline & Beecham
Pharmaceuticals
King of Prussia, Pennsylvania

Jerome M. Bailey
Hewlett-Packard Company
Palo Alto, California

Andreas D. Baxevanis
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Alain Bernard
Glaxo Institute for Molecular Biology
Geneva, Switzerland

Mark S. Boguski
National Center for Biotechnology
Information
Bethesda, Maryland

Jeffrey C. Boyington
National Institute of Allergy &
Infectious Diseases
Rockville, Maryland

Ineke Braakman
University of Amsterdam Academic
Medical Centre
The Netherlands

Anne V. Broadhurst
Roche Products Ltd.
Welwyn Garden City
United Kingdom

Crawford Brown
British Bio-Technology
Cowley, United Kingdom

Daniel C. Bundy
Eastman Kodak Company
New Haven, Connecticut

Stephen Buratowski
Harvard Medical School
Boston, Massachusetts

Dean Burgi
Genomyx
Foster City, California

Steven A. Carr
SmithKline & Beecham
Pharmaceuticals
King of Prussia, Pennsylvania

Nigel Carter
The Salk Institute
La Jolla, California

J. David Castle
University of Virginia Medical School
Charlottesville, Virginia

Su Chen
Chiron Corporation
Emeryville, California

Christopher C.Q. Chin
University of Texas Medical School
Houston, Texas

Lewis A. Chodosh
University of Pennsylvania
Philadelphia, Pennsylvania

Jeffrey J. Clare
Wellcome Research Laboratories
Beckenham, United Kingdom

John W. Crabb
W. Alton Jones Cell Science Center
Lake Placid, New York

Mark W. Crankshaw
Washington University School of
Medicine
St. Louis, Missouri

Dan L. Crimmins
Washington University School of
Medicine
St. Louis, Missouri

David R. Davies
National Institute of Diabetes and
Digestive and Kidney Diseases
Bethesda, Maryland

Nancy D. Denslow
University of Florida
Gainesville, Florida

Tara M. DeSilva
The Wistar Institute
Philadelphia, Pennsylvania

Sandra Diaz
University of California San Diego
La Jolla, California

Michael P. DiGiovanna
Yale University School of Medicine
New Haven, Connecticut

W. Scott Dodson
W. Alton Jones Cell Science Center
Lake Placid, New York

Tamara L. Doering
Cornell University Medical College
New York, New York

Ben M. Dunn
University of Florida
Gainesville, Florida

Paul T. Englund
Johns Hopkins Medical School
Baltimore, Maryland

Joseph Fernandez
The Rockefeller University
New York, New York

Gregg B. Fields
University of Minnesota
Minneapolis, Minnesota

Elizabeth Fowler
AutoImmune, Inc.
Lexington, Massachusetts

Hudson H. Freeze
LaJolla Cancer Research Foundation
LaJolla, California

Sean R. Gallagher
Motorola Corporation
Tempe, Arizona

Robert Garian
George Mason University
Fairfax, Virginia

Kieran F. Geoghegan
Pfizer, Inc.
Groton, Connecticut

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

Warren Gish
Washington University School of Medicine
St. Louis, Missouri

David P. Goldenberg
University of Utah
Salt Lake City, Utah

Gregory A. Grant
Washington University School of Medicine
St. Louis, Missouri

David Gray
Chiron Corporation
Emeryville, California

Jean Gruenberg
University of Geneva Sciences II
Geneva, Switzerland

Lars Hagel
Pharmacia Biotech AB
Uppsala, Sweden

Stephen E. Harding
University of Nottingham School of Biology
Sutton Bonington, United Kingdom

Sandra Harper
The Wistar Institute
Philadelphia, Pennsylvania

Gerald W. Hart
University of Alabama
Birmingham, Alabama

Daniel N. Hebert
Yale University School of Medicine
New Haven, Connecticut

William J. Henzel
Genentech, Inc.
So. San Francisco, California

Alison B. Hickman
National Institute of Diabetes and Digestive and Kidney Diseases
Bethesda, Maryland

David R. Higgins
Invitrogen Corporation
San Diego, California

Bret Holmquist
University of California
Riverside, California

Jeffrey D. Hulmes
W. Alton Jones Cell Science Center
Lake Placid, New York

Caroline S. Jackson
National Institute for Medical Research
London, United Kingdom

Kornelia Jumel
University of Nottingham School of Biology
Sutton Bonington, United Kingdom

James T. Kadonaga
University of California San Diego
La Jolla, California

Robert M. Kennedy
Pharmacia Biotech, Inc.
Piscataway, New Jersey

Leslie A. Kerrigan
Osiris Therapeutics
Baltimore, Maryland

Radha G. Krishna
University of Texas Medical School
Houston, Texas

Stanley R. Krystek, Jr.
Bristol-Myers Squibb Pharmaceutical Research Institute
Princeton, New Jersey

Thomas M. Laue
University of New Hampshire
Durham, New Hampshire

Shu-Mei Liang
North American Vaccine Corp.
Beltsville, Maryland

Rex Lovrien
University of Minnesota
St. Paul, Minnesota

Elizabeth J. Luna
Worcester Foundation for Biomedical Research
Shrewsbury, Massachusetts

George Lunn
Baltimore, Maryland

Johnny Ma
Chiron Corporation
Emeryville, California

Henryk Mach
Merck Research Laboratories
West Point, Pennsylvania

Anthony I. Magee
National Institute for Medical Research
London, United Kingdom

Adriana E. Manzi
University of California San Diego
La Jolla, California

George I. Makhatadze
Texas Tech University
Lubbock, Texas

Ronen Marmorstein
The Wistar Institute
Philadelphia, Pennsylvania

Daumantas Matulis
University of Minnesota
St. Paul, Minnesota

John McCoy
Genetics Institute
Cambridge, Massachusetts

Scott Medberry
Amersham Pharmacia Biotech
San Francisco, California

Jill Meisenhelder
The Salk Institute
LaJolla, California

William J. Metzler
Bristol-Myers Squibb Pharmaceutical Research Institute
Princeton, New Jersey

George Michaels
George Mason University
Fairfax, Virginia

C. Russell Middaugh
Merck Research Laboratories
West Point, Pennsylvania

Chad G. Miller
Hewlett-Packard Company
Palo Alto, California

Sheenah M. Mische
The Rockefeller University
New York, New York

Mary Moyer
Glaxo Research Institute
Research Triangle Park, North Carolina

Jacek Mozdzanowski
The Wistar Institute
Philadelphia, Pennsylvania

SCIANI / CENTRAL UNIVERSITY LIBRARY

Cheryl Isaac Murphy
Cambridge Biotech Corporation
Worcester, Massachusetts

Jiri Novotny
Bristol-Myers Squibb Pharmaceutical Research Institute
Princeton, New Jersey

B.F. Francis Ouellette
National Center for Biotechnology Information
Bethesda, Maryland

Roger H. Pain
Jozef Stefan Institute
Ljubljana, Slovenia

Ira Palmer
National Institutes of Health
Bethesda, Maryland

Mark Payton
Glaxo Institute for Molecular Biology
Geneva, Switzerland

Kevin J. Petty
University of Texas Southwestern Medical Center
Dallas, Texas

Mary C. Phelan
Greenwood Genetic Center
Greenwood, South Carolina

Allen T. Phillips
Pennsylvania State University
University Park, Pennsylvania

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

Hidde L. Ploegh
Massachusetts Institute of Technology
Cambridge, Massachusetts

Leland D. Powell
University of California San Diego
La Jolla, California

Kathryn R. Radford
Glaxo Institute for Molecular Biology
Geneva, Switzerland

Hervé E. Recipon
National Center for Biotechnology Information
Bethesda, Maryland

David F. Reim
The Wistar Institute
Philadelphia, Pennsylvania

Pier Giorgio Righetti
University of Milan
Milan, Italy

Lise R. Riviere
Pfizer, Inc.
Groton, Connecticut

Stuart J. Rodda
Chiron Technologies Pty. Ltd.
Victoria, Australia

Michael A. Romanos
Wellcome Research Laboratories
Beckenham, United Kingdom

Keith Rose
University Medical Center
Geneva, Switzerland

Robert R. Roussel
Dartmouth College
Hanover, New Hampshire

Gregory D. Schuler
National Center for Biotechnology Information
Bethesda, Maryland

Ton N.M. Schumacher
Massachusetts Institute of Technology
Cambridge, Massachusetts

R. K. Scopes
LaTrobe University
Bundoora, Australia

William H. Scouten
Utah State University
Logan, Utah

Brian Seed
Massachusetts General Hospital and Harvard Medical School
Boston, Massachusetts

Bartholomew M. Sefton
The Salk Institute
La Jolla, California

Kentaro Semba
University of Tokyo
Tokyo, Japan

Shirish Shenolikar
Duke University Medical Center
Durham, North Carolina

Michael H. Simonian
Beckman Instruments, Inc.
Fullerton, California

Alan J. Smith
Stanford University Medical Center
Stanford, California

David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

Timothy A. Springer
Center for Blood Research
Harvard Medical School
Boston, Massachusetts

Earle Stellwagen
University of Iowa
Iowa City, Iowa

David F. Stern
Yale University School of Medicine
New Haven, Connecticut

Alasdair C. Steven
National Institute of Arthritis and Musculoskeletal and Skin Diseases
Bethesda, Maryland

Kathryn L. Stone
Yale University
New Haven, Connecticut

Robert L. Strausberg
National Institutes of Health
Bethesda, Maryland

Susan L. Strausberg
University of Maryland
Rockville, Maryland

John T. Stults
Genentech, Inc.
So. San Francisco, California

Shyamsundar Subramanian
Chiron Corporation
Emeryville, California

Peter D. Sun
National Institute of Allergy & Infectious Diseases
Rockville, Maryland

Paul Tempst
Memorial Sloan-Kettering Cancer Center
New York, New York

BUJASI / CENTRAL UNIVERSITY LIBRARY

Theodore J. Tsomides
Massachusetts Institute of Technology
Cambridge, Massachusetts

Jeanine A. Ursitti
The Wistar Institute
Philadelphia, Pennsylvania

Ajit Varki
University of California San Diego
La Jolla, California

Karen A. West
W. Alton Jones Cell Science Center
Lake Placid, New York

Alan Williams
Pharmacia Biotech, Inc.
Piscataway, New Jersey

Kenneth R. Williams
Yale University
New Haven, Connecticut

Paul T. Wingfield
National Institutes of Health
Bethesda, Maryland

Finn Wold
University of Texas Medical School
Houston, Texas

Paula Wolf
Massachusetts Institute of Technology
Cambridge, Massachusetts

Louis Zumstein
Introgen Therapeutics, Inc.
Houston, Texas

BCU IASI / CENTRAL UNIVERSITY LIBRARY

1 STRATEGIES OF PROTEIN PURIFICATION AND CHARACTERIZATION

1.1	**Overview of Protein Purification and Characterization**	**1.1.1**
	Aims and Objectives	1.1.1
	Sources of Material for Protein Purification	1.1.2
	Detection and Assay of Proteins	1.1.3
	Methods for Separation and Purification of Proteins	1.1.3
	Characterization of the Protein Product	1.1.4
	The Protein Purification Laboratory	1.1.5
1.2	**Strategies for Protein Purification**	**1.2.1**
	Classification of Proteins	1.2.1
	Strategies for Protein Purification	1.2.1
1.3	**Protein Purification Flow Charts**	**1.3.1**
	Soluble Recombinant Proteins	1.3.1
	Insoluble Recombinant Proteins	1.3.1
	Soluble Nonrecombinant Proteins	1.3.3
	Membrane-Associated and Insoluble Nonrecombinant Proteins	1.3.5

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Copyright © 1995 by John Wiley & Sons, Inc.

Overview of Protein Purification and Characterization

AIMS AND OBJECTIVES

Protein purification has an over 200-year history: the first attempts at isolating substances from plants having similar properties to "egg albumen," or egg white, were reported in 1789 by Fourcroy. Many proteins from plants were purified in the nineteenth century, though most would not be considered pure by modern standards. A century later, ovalbumin was the first crystalline protein obtained (by Hofmeister in 1889). The year 1989 may not go down in history as a milestone in protein chemistry, but since then there has been a resurgence of interest in proteins after more than a decade of gene excitement.

The aims of protein purification, up until the 1940s, were simply academic. To then, even the basic facts of protein structure were not fully appreciated, and pure proteins were needed just to study structure and test the rival theories of the pre-DNA days. During the Second World War, an acute need for blood proteins led to development of the Cohn fractionation procedure for purification of albumin and other proteins from serum (Cohn et al., 1946). This was the inception of large-scale protein purifications for commercial purposes; Cohn fractionation continues to be used to this day.

As more proteins, and particularly enzymes, were purified and crystallized, they started to be used increasingly in diagnostic assays and enzymatic analyses, as well as in the large-scale food, tanning, and detergent industries. Many enzymes used in industry are not in fact very pure, but as long as they do the job, that is sufficient. "Process" enzymes such as α-amylase, proteases, and lipases are pro-duced in ton quantities, mainly as secretion products in bacterial cultures, and may undergo only limited purification processes to mini-mize costs. At the other extreme, enzyme products for research and analysis require a high degree of purification to ensure that contaminating activities do not interfere with the intended use. Anyone familiar with molecular biology enzymes will appreciate how minute levels of contamination of DNase or RNase can completely destroy carefully planned experiments.

The 1960s and 1970s could be described as the peak years for protein and enzyme research, and most of the methods used in protein purification were established by then, at least in their principles. More recent developments have been mainly in instrumentation designed to optimize the application of each methodology. Developments in instrumentation have been stimulated by the rapid progress in molecular biology, because gene isolation has often been preceded by isolation of the gene product. Because such products can now be characterized sufficiently (i.e., partially sequenced) using minute amounts of protein, the need for large-scale or even moderate-scale procedures has decreased. Hence there has been an explosive development of modern equipment designed specifically for dealing with amounts of protein in the milligram to microgram range. On the other hand, structural studies using X-ray crystallography and nuclear magnetic resonance (NMR) require hundreds of milligrams of pure protein, so larger-scale equipment and procedures are still needed in the research laboratory.

The nature of the proteins studied has also changed substantially. Whereas enzymes were once the most favored subjects, they have now been superceded by nonenzymatic proteins such as growth factors, hormone receptors, viral antigens, and membrane transporters. Many of these occur in minute amounts in the natural source, and their purification can be a major task. Heroic efforts in the past have used kilogram quantities of rather unpleasant starting materials, such as human organs, and ended up with a few micrograms of pure product. It is now more usual, however, to take the genetic approach: clone the gene before the protein has been isolated or even properly identified, then express it in a suitable host cell culture or organism. The expression level may be orders of magnitude higher than in the original source, which will make purification a relatively simple task. It can be useful to know beforehand some physical properties of the protein, to facilitate the development of a suitable purification protocol from the recombinant source. On the other hand, there are now several ways of preparing fusion proteins, which can be purified by affinity techniques without any knowledge of the properties of the target protein. Moreover, there are ways of modifying the expressed product to simplify purification further.

Copyright © 1995 by John Wiley & Sons, Inc.

LIBRARY

Thus the approach to protein purification must first take into account the reason it is being done, as the methods will vary greatly with different requirements. At one extreme is the one-of-a-kind purification, in a well-financed and equipped laboratory, that is carried out to obtain a small amount of product for sequencing so that gene isolation can proceed. In this case, expense of equipment and reagents may be no problem, and a very low overall recovery of product can be acceptable, provided it is pure enough. At the other extreme are the requirements of commercial production of a protein in large amounts on a continuing basis, where high recovery and economy of processing are the chief parameters to be considered. There are many intermediate situations as well.

Many publications in the area of protein research are entitled "Purification and characterization of...," and describe a purification procedure in sufficient detail that it can be reproduced in another laboratory. The characterization section may include structural, functional, and genetic information, and carrying out such studies is likely to require at least milligram quantities of pure protein. Ideally the purification should involve a small number of steps, with good recovery at each step. If the recovery is poor (<50% at any step), however, there should be some indication of what happened to the missing activity. Has it been discarded in the other fractions for the sake of purity, or does it represent a true loss of activity? If the latter, then the end-product may be less than fully active despite apparent homogeneity indicated by standard analysis. The choice between recovery and purification at each step can be problematical; taking a narrow cut of a chromatographic peak may provide a very pure fraction, at the expense of losing a good deal of less pure active component on either side. In making such decisions, the objective of the exercise must be kept in mind: if yield is not important, then the choice of poor yield for the sake of purity may be logical.

By far the most important requirement of a publication is reproducibility of the method reported. It is not sufficient to have carried out the process only once if it is expected that other investigators will want to repeat it. There are always factors that influence the process that may be overlooked at first, and which if varied slightly can have a major effect on the purification procedure. The reported process should always be repeated exactly as described before submitting the manuscript for publication. There is one exception, namely, the case where purification was conducted simply to obtain enough protein for sequencing and gene isolation; if that was achieved, there should be no need to provide instructions for repetition.

SOURCES OF MATERIAL FOR PROTEIN PURIFICATION

For many people embarking on a protein purification project, there is no choice of material. They are studying a particular biological tissue or organism, and the objective is to purify a protein from that source. However, there may be approaches that can make the project simpler. If, for instance, the source is difficult to obtain in large amounts, it may be best to carry out at least preliminary trials on a source species more readily obtained. The most obvious and relevant example is when the species being studied is *Homo sapiens*, and tissue samples are not readily available for practical or ethical reasons, or both. In this case, it is usual to go to where mammalian tissue is readily available (i.e., an abattoir) and work with bovine, ovine, or porcine sources. Alternatively, if quantity of tissue is not a problem, the humble laboratory rat may suffice. Once a protocol for purifying the protein from substitute sources has been worked out, it will be much easier to develop one using human material—the identical procedure may work satisfactorily. Proteins differ to a fairly small extent between species that have diverged within about 100 million years, a time frame that groups together most higher mammals. Thus the behavior of proteins derived from different animals with respect to the various fractionation procedures is likely to be similar, and a protocol worked out for pig tissues is likely to need only minor adjustments for application to human tissues.

A second example is where the interest is mainly on the function of a protein, especially an enzyme, for which functions and actions have generally been strongly conserved through evolution. In that case, a preliminary screening of potential sources, or, better still, the literature, should come up with a raw material that is best suited to the investigator's purposes. Considerations should include the following: (1) What functions are required of the end product? For instance, an enzyme having a low K_m may be needed, so selecting the source with the highest activity may not suffice. (2) How convenient is it to grow or obtain the raw material, and are there problems concerning pathogenicity or extractability? (3) Does the quantity of the protein vary with growth conditions or age, and does it deteriorate in situ

if left too long? Obviously one requires a source that reliably produces the highest amount of the desired protein per unit volume to maximize the chances of developing a good purification procedure. (4) What storage conditions are required for the raw material? It is important to consider that fresh raw material may not be immediately available whenever a purification is attempted.

The above considerations are relevant to the traditional situation for commencing a protein purification project. It is becoming increasingly common, however, for proteins to be purified as recombinant products using techniques in which the gene is expressed in a host organism or in cultured cells. This of course requires the gene encoding the protein of interest to be available. Until the mid-1980s, such material was usually obtained by hybridization of an oligonucleotide synthesized according to amino acid sequence information. This required the protein to have been purified first, so the initial task of protein purification still needed to be done at least once. More recently, genetic techniques have permitted the isolation of many genes encoding known proteins, even though the proteins may never have been studied directly. Moreover, with the expansion of the Human Genome Project and related DNA sequencing efforts, many genes for both known and unknown proteins will become available and will be able to be expressed in recombinant form without ever being purified from the host species. As a result some completely new considerations for protein purification come into play, including the possibility of modifying the gene structure not only to increase expression level and alter the protein product itself to enhance a desired function, but equally importantly to aid in purification. Recombinant proteins may be expressed in bacteria, yeasts, insect cells, and animal tissue cultures. Further details may be found in *UNIT 1.2*.

DETECTION AND ASSAY OF PROTEINS

During a protein purification procedure there are two measurements that need to be made, preferably for each fraction. Measurements both of the total protein and of the amount (usually bioactivity) of the desired protein must be made. Details of the most commonly used assay methods are given in Chapter 3. It is not possible to isolate a protein without a method of determining whether it is present; an assay, either quantitative or at least semi-quantitative, indicating which fraction contains the most of the desired protein is essential.

Assays may range from the quick-and-easy type (e.g., instantaneous spectrophotometric measurement of enzyme activity) to long and tedious bioassays that may take days to produce an answer. The latter situation is very difficult, because by the time one knows where the protein is, it may be "was," owing to degradation or inactivation. Moreover, this may not become clear until the next step has been completed and its products assayed. Any assay that is quick is therefore advantageous, even if it means a sacrifice of accuracy for speed.

Measurement of total protein is useful, as it indicates the degree of purification at each step. However, unless the next step critically depends on how much protein is present, total protein measurement is not extremely important: a small sample can be put aside and measured later, when the purification is complete. It is, however, very important to know how much protein is present in the final, presumed pure sample, as this will indicate the specific activity (if the protein has an activity), which can be compared with other preparations. The general object is to obtain as high a specific activity as possible (taking into account recovery considerations), which means retaining as much of the desired protein as possible while ending up with as little total protein as possible.

METHODS FOR SEPARATION AND PURIFICATION OF PROTEINS

The methods available for protein purification range from simple precipitation procedures used since the nineteenth century to sophisticated chromatographic and affinity techniques that are constantly undergoing development and improvement. Methods can be classified in several alternative ways—perhaps one of the best is based on the properties of the proteins that are being exploited. Thus the methods can be divided into four distinct but interrelated groups depending on protein characteristics: surface features, size and shape, net charge, and bioproperties.

Methods Based on Surface Features of Proteins

Surface features include charge distribution and accessibility, surface distribution of hydrophobic amino acid side chains, and, to a lesser extent, net charge at a given pH (see discussion of net charge). Methods exploiting surface fea-

BOULDER COLORADO STATE LIBRARY

tures mainly depend on solubility properties. Differences in solubility result in precipitation by various manipulations of the solvent in which the proteins are solubilized. Methods for obtaining an extract containing the desired protein in soluble form are given in Chapter 4. The solvent, nearly always water containing a low concentration of buffer salts, can be treated to alter properties such as ionic strength, dielectric constant, pH, temperature, and detergent content, any of which may selectively precipitate some of the proteins present. Conversely, proteins may be selectively solubilized from an insoluble state by manipulation of the solvent composition. The surface distribution of hydrophobic residues is an important determinant of solubility properties; it is also exploited in hydrophobic chromatography, both in the reversed-phase mode (*UNIT 11.6*) and in aqueous-phase hydrophobic-interaction chromatography (*UNIT 8.4*).

Also included in this group is the highly specific technique of immunoaffinity chromatography, in which an antibody directed against an epitope on the protein surface is used to pull out the desired protein from a mixture.

Methods Based on Whole Structure: Size and Shape

Although the size and shape of proteins can have some influence on solubility properties, the chief method of exploiting these properties is gel-filtration chromatography (*UNIT 8.3*). In addition, preparative gel electrophoresis makes use of differences in molecular size. Proteins range in size from the smallest classified as proteins rather than polypeptides, around 5000 Da, up to macromolecular complexes of many million daltons. Many proteins in the bioactive state are oligomers of more than one polypeptide (see *UNIT 1.2*), and these can be dissociated, though normally with loss of overall structure. Thus many proteins have two "sizes": that of the native state, and that (or those) of the polypeptides in the denatured and dissociated state. Gel-filtration procedures normally deal only with native proteins, whereas electrophoretic procedures commonly involve separation of dissociated and denatured polypeptides.

Methods Based on Net Charge

The two techniques that exploit the overall charge of proteins are ion-exchange chromatography (by far the most important) and electrophoresis (Chapter 10). Ion exchangers bind charged molecules, and there are essentially only two types of ion exchangers, anion and cation. The net charge of a protein depends on the pH—positive at very low pH, negative at high pH, and zero at some specific point in between, termed the isoelectric point (pI). It should be stressed that at the pI a protein has a great many charges; it just happens that at this pH the total negatives exactly equal the total positives. The most charged state (disregarding the charge sign) is in the pH range 6.0 to 9.0. This is the most stable pH range for most proteins, as it encompasses common physiological pH values. Ion exchangers consist of immobilized charged groups and attract oppositely charged proteins. They provide the mode of separation that has the highest resolution for native proteins. High-performance reversed-phase chromatography has equivalent or even better resolution, but it generally involves at least partial denaturation during adsorption and so is not recommended for sensitive proteins such as enzymes. Protein purification using ion-exchange chromatography has mainly employed positively charged anion exchangers, for the simple reason that the majority of proteins at neutral pH are negatively charged (i.e., have a low isoelectric point). Details of methodology are found in *UNIT 8.2*.

Methods Based on Bioproperties (Affinity)

A powerful method for separating the desired protein from others is to use a biospecific method in which the particular biological property of the protein is exploited. The affinity approach is limited to proteins that have a specific binding property, except that proteins are theoretically able to be purified by immunoaffinity chromatography (*UNIT 9.1*), which is the most specific of all affinity techniques. Most proteins of interest do have a specific ligand: enzymes have substrates and cofactors, and hormone-binding proteins and receptor molecules are designed to bind specifically and tightly to particular hormones and other factors. Immobilization of the ligand to which the protein binds (or of antibody to the protein) enables selective adsorption of the desired protein in the technique known as affinity chromatography (Chapter 9). There are also non-chromatographic modes of exploiting biospecific interactions.

CHARACTERIZATION OF THE PROTEIN PRODUCT

Once a pure protein is obtained, it may be employed for a specific purpose, such as enzymatic analysis (e.g., glucose oxidase and lac-

PGU IASRI LIBRARY

tate dehydrogenase), or as a therapeutic agent (e.g., insulin and growth hormone). However, it is normal, when a protein has been isolated for the first time, to characterize it in terms of structure and function. Several features are generally expected in characterization of a new protein. These include molecular weight, or at least the size of the subunit(s), determined by SDS-polyacrylamide gel electrophoresis (Chapter 10) and/or gel filtration (*UNIT 8.3*). Spectral properties such as the UV spectrum (Trp and Tyr content), circular dichroism (CD) spectrum (secondary structure), and special characteristics of proteins with prosthetic groups (e.g., quantitation and spectra) may be presented. The quantity and nature of carbohydrates on glycoproteins should be determined (Chapter 12). Also, if the gene has not already been reported, some amino-terminal sequence analysis should be given, if at all possible, along with the results of a database search for similar sequences (*UNIT2.1*). Functional proteins should be demonstrated to have the appropriate function, and for enzymes detailed kinetic characterization is appropriate. Ultimately the full three-dimensional structure of the protein may be determined, which will require crystals: any successful crystallization attempts should be reported.

THE PROTEIN PURIFICATION LABORATORY

The requirements for a protein purification laboratory cannot be exactly formulated because they depend greatly on the types and amounts of proteins being isolated. To cover all eventualities, it would be necessary to have one set of equipment to deal with submicrogram quantities and another set to deal with multigram quantities—a range of around 10^8! One laboratory dedicated to protein purification may not need small-scale equipment if, for example, it works with plasma proteins that are always available in large quantities. Another may have all the latest in high-performance equipment but not be able (nor need) to handle quantities of protein in excess of a few milligrams.

If it is assumed that neither extreme in quantity is to be attempted, and that the laboratory is handling a variety of protein types and sources, then certain basic pieces of equipment are needed. Obtaining the starting material and making an extract of it require homogenization equipment and centrifuges to remove insoluble residues. Preliminary fractionation, when starting with a crude extract of tissue or cells, requires equipment and materials that will not become clogged by particulates. Adsorbents and similar materials used at the first step should be relatively inexpensive so that when performance falls off after a few uses, owing to intransigent impurity buildup, they can be discarded. It is also relevant that a larger amount is handled at the initial step than later steps; therefore, reagent expense can be an important consideration. After the first one or two steps, the sample should be sufficiently clean and clear to enable use of high-performance equipment.

High-performance liquid chromatography, or HPLC, is a term with a variety of meanings. To some it refers exclusively to reversed-phase chromatography; to others it includes all sorts of chromatography provided that the equipment is fully automated and high-performance adsorbents are used. A high-performance system designed specifically for proteins—introduced by Pharmacia Biotech (see *SUPPLIERS APPENDIX*) called Fast Protein Liquid Chromatography, or FPLC—uses standard protein chromatographies such as ion exchange, hydrophobic interaction, and gel filtration. Scaleup is possible with larger equipment based on the FPLC design, so that laboratory development can be quickly translated to large-scale production. FPLC is designed to separate proteins in their native active configuration, whereas reversed-phase HPLC often causes at least transient denaturation during adsorption and elution. Reversed-phase HPLC has a high resolving power, but it is best suited to peptides and proteins smaller than ~30 kDa. Chromatography run with older-style low-pressure adsorbents is sometimes referred to as "low-performance" or "open-column" chromatography; neither of those descriptions is necessarily accurate. Simple fraction collector and monitoring equipment is needed. This equipment will be used for larger-scale operations (tens of milligrams of protein and upward), probably at an earlier stage in the protocol than with HPLC.

Various columns, both prepacked with proprietary adsorbents and empty for self-packing, will be needed, with the sizes and types depending on the scale of operations. Several anion-exchange columns (different sizes), one or two cation-exchange columns, and gel-filtration media are essential, along with a range of alternative adsorbents such as hydrophobic interaction materials, dyes, hydroxyapatite, and chromatofocusing and specialist affinity media.

MASS CENTRAL LIBRARY

Fully equipped protein purification laboratories should also have preparative electrophoresis and isoelectric focusing apparatuses for rare occasions when other techniques fail to give sufficient separation.

In addition to equipment used in the actual fractionation processes, a variety of other items are needed. In particular it should be possible to change buffers quickly and to concentrate protein solutions with ease. These operations require such things as dialysis membranes (*APPENDIX 3B*), ultrafiltration cells, and gel-exclusion columns of various sizes (*UNIT 8.3*).

Finally, equipment for assaying and analyzing the preparations is needed. Most such equipment is fairly standard in biochemical laboratories and includes spectrophotometers, scintillation counters, analytical gel and capillary electrophoresis apparatuses, immunoblotting materials, and immunochemical reagents. A listing of standard equipment is found in *APPENDIX 2D*.

Literature Cited

Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.J., Ashworth, J.N., Melin, M., and Taylor, H.L. 1946. Preparation and properties of serum and plasma proteins. IV. A system for the separation into fractions of the proteins and lipoprotein components of biological tissues and fluids. *J. Am. Chem. Soc.* 68:459-475.

Key References

Deutscher, M.P. (ed.) 1990. Guide to protein purification. *Methods Enzymol.* 182:1-894.

Extensive collection of purification methods with some general protocols and examples.

Janson, J.-C. and Ryden, L.G. 1989. Protein Purification: Principles, High Resolution Methods, and Applications. VCH Publishers, New York.

A useful collection of methods and examples.

Kennedy, J.F. and Cabral, J.M. (eds.) 1993. Recovery Processes for Biological Materials. John Wiley & Sons, New York.

A useful introduction to the problems of large-scale methods.

Kenny, A. and Fowell, S. (eds.) 1992. Practical protein chromatography. *Methods Mol. Biol.* 11:1-327.

Extensive descriptions of affinity chromatographic techniques with protocols and recipes.

Scopes, R.K. 1993. Protein Purification, Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.

General principles of all the main techniques used in purifying proteins. A useful laboratory handbook; does not include recipes or procedures for specific proteins.

Contributed by R. K. Scopes
La Trobe University
Bundoora, Australia

PCU IAST/ CENTRAL UNIVERSITY LIBRARY

Strategies for Protein Purification

CLASSIFICATION OF PROTEINS

As with most heterogeneous collections of things, proteins can be classified in several different ways, such as by function, by structure, or by physicochemical characteristics. Each protein species consists of identical molecules with exactly the same size, amino acid sequence, and three-dimensional shape. In this way a solution of a mixture of proteins differs from a solution of synthetic polymers or sheared DNA, both of which contain a complete spectrum of possible sizes centered around the average. The protein mixture has only discrete sizes of molecules corresponding to each type of protein present. Although we could classify proteins by size, it would be of limited use, as there is usually no obvious relationship between size and function.

A more useful structural classification takes into consideration shape and oligomeric structure (Table 1.2.1). In part, structure reflects biological location and origin. Simple, fairly rigid protein molecules occur in the extracellular environment, more complex and readily deactivated molecules are found intracellularly, and hydrophobic proteins are associated with membranes.

Classification by function is more relevant (Table 1.2.2). Proteins can be simply stores of amino acids, can be structural, or can have specific binding functions. The most "functional" proteins are enzymes, which have both binding and catalytic roles. In part, this reflects the degree to which the detailed structure is a requirement for the protein's function, which in turn relates to conservation of structure through evolution. But as with every attempt at classification, there are going to be examples that do not fit the pattern well. Most proteins of interest to the pharmaceutical industry belong to the general class of binding proteins, for instance, hormones [e.g., insulin and bovine somatostatin (BST)], viral antigens (e.g., hepatitis B antigen), growth factors [e.g., interferons, interleukins and colony-stimulating factors (CSFs)], and antibodies.

STRATEGIES FOR PROTEIN PURIFICATION

Soluble Extracellular Proteins

The source of soluble extracellular proteins is the extracellular medium, whether it be an animal source such as blood or spinal fluid, or a culture medium in which bacterial, fungal, animal, or plant cultures have been grown. Generally these do not contain a large number of different proteins (blood is an exception), and the desired protein may be a major component, especially if produced as the result of recombinant expression. Nonetheless, the protein in the starting material may be quite dilute,

Table 1.2.1 Classification of Proteins by Structural Characteristics

Structural characteristic	Examples	Comments
Monomeric	Lysozyme, growth hormone	Usually extracellular; often have disulfide bonds
Oligomeric		
Identical subunits	Glyceraldehyde-3-phosphate dehydrogenase, catalase, alcohol dehydrogenase, hexokinase	Mostly intracellular enzymes; rarely have disulfide bonds
Mixed subunits	Aspartate carbamoyltransferase, pertussis toxin	Allosteric enzymes; different subunits have separate functions
Membrane-bound		
Peripheral	Mitochondrial ATPase, alkaline phosphatase	Readily solubilized by detergents
Integral	Porins, cytochromes P_{450}, insulin receptor	Require lipid for stability
Conjugated	Glycoproteins, lipoproteins, nucleoproteins	Many extracellular proteins contain carbohydrate

Copyright © 1995 by John Wiley & Sons, Inc.

BOU IASL CENTRAL UNIVERSITY LIBRARY

Table 1.2.2 Classification of Proteins by Function

Function	Examples
Amino acid storage	Seed proteins (e.g., gluten), milk proteins (e.g., casein)
Structural	
Inert	Collagen, keratin
With activity	Actin, myosin, tubulin
Binding	
Soluble	Albumin, hemoglobin, hormones
Insoluble	Surface receptors (e.g., insulin receptor), antigens (e.g., viral coat proteins)
With activity	Enzymes, membrane transporters (e.g., amino acid uptake systems, ion pumps)

and a large volume may therefore need to be processed. The starting fluid may also contain many compounds other than proteins, whose behavior must be taken into account. The first stage should aim mainly to reduce the volume and get rid of as much nonprotein material as possible; some protein-protein separation is also useful, but not essential. No general rules can be given, but a batch adsorption method using an inexpensive material such as hydroxyapatite, ion-exchange resin, immobilized metal affinity chromatography (IMAC) medium, or affinity adsorbent is best, if feasible. Following the first step, the sample should be in a form that is amenable to standard purification processes such as precipitation and column chromatography (Chapter 8).

Intracellular (Cytoplasmic) Proteins

To obtain soluble intracellular proteins (which are mainly enzymes), cells must be broken open or lysed to release their soluble contents. The ease with which cell disruption can be accomplished varies considerably; animal cells are readily broken, as are many bacteria, but plants and fungi have tough cell walls. Methods for obtaining cell extracts are given in Chapter 4. The macromolecular soluble contents of cells are mainly proteins, with nucleic acids as a minor but significant component. Bacterial extracts may be viscous unless DNase is added to break down the long DNA molecules. Although chromatographic procedures can be applied to crude extracts, valuable high-performance materials should not be employed in the first step, as there are always compounds, including unstable proteins, that may bind to them and be difficult to remove.

Membrane-Associated Proteins

There are two approaches to isolating a membrane-associated protein. In one method, the relevant membrane fraction can first be prepared and then used to isolate the protein. Alternatively, whole tissue can be subjected to an extraction that solubilizes the membranes and releases the cytoplasmic contents as well. The former is much better in that purification is accomplished by isolating the membranes: the specific activity of the solubilized membrane fraction will be much higher than in the second method. However, the process of purifying the membrane fraction may lead to substantial losses, and it may be difficult to scale up. If total recovery of the protein is more important than purity, a whole-tissue extract is likely to be more appropriate. Although this means that a greater degree of purification is needed, the fact that membrane proteins have, by definition, properties somewhat different from those of cytoplasmic proteins permits some very effective purification steps (e.g., hydrophobic chromatography or fractional solubility separation).

Peripheral membrane proteins are only loosely attached and may be released by gentle conditions such as high pH, EDTA, or low (nonionic) detergent concentrations. Once in solution, some peripheral proteins no longer require the presence of detergent to maintain their solubility. Integral membrane proteins are much more difficult—they require high concentrations of detergent for solubilization (i.e., complete solubilization of the membrane is needed to release them) and generally are neither soluble nor stable in the absence of detergent. It is sometimes necessary to maintain

BCU IAS CENTRAL UNIVERSITY LIBRARY

natural phospholipids in association with the proteins in order to maintain activity. Even though the final objective may not require activity (e.g., amino acid sequencing), there is a need for some sort of assay during the purification process to determine where the protein is. If a particular band on a gel is known to be the desired protein, then no other assay is needed and loss of bioactivity can be allowed.

Purification processes may be affected by the presence of detergents. The problem of association with detergent micelles makes purifying integral membrane proteins difficult; the close association of the different proteins originating from membranes often results in very poor separation in conventional fractionation procedures.

Insoluble Proteins

Natural proteins that are insoluble in normal solvents are generally structural proteins, which are sometimes cross-linked by post-translational modification. The first stage of purification is obvious—it involves extracting and washing away all proteins that are soluble, leaving the residue containing the desired material. Further purification in a native state, however, may be impossible; extracting away other proteins using more vigorous solvents or attempting to solubilize the target protein may destroy the natural structure. Cross-linked proteins such as elastin or aged collagen cannot be dissolved without breaking the cross-links, and the individual proteins may even be cross-linked together.

Insoluble Recombinant Proteins (Inclusion Bodies)

A major new class of insoluble proteins are recombinant proteins expressed (usually in *Escherichia coli*) as inclusion bodies. These are dense aggregates found inside cells that consist mainly of a desired recombinant product, but in a nonnative state. Inclusion bodies may form for a variety of reasons, such as insolubility of the product at the concentrations being produced, inability to fold correctly in the bacterial environment, or inability to form correct, or any, disulfide bonds in the reducing intracellular environment. Their purification is simple, since the inclusion bodies can be separated by differential centrifugation from other cellular constituents, giving almost pure product; the problem is that the protein is not in a native state, and is insoluble. Some methods for obtaining an active product from inclusion bodies are described in *UNIT 6.3*.

Soluble Recombinant Proteins

Recombinant proteins that are not expressed in inclusion bodies either will be soluble inside the cell or, if using an excretion vector, will be extracellular (or, if *E. coli* is the host, possibly periplasmic). They can be purified by conventional means. In some systems, expression is so good that the desired product is the major protein present and its purification is relatively simple. In systems where the expression level is low, the purification process can be tedious, though easier, it is hoped, than isolation from the natural source. It should be remembered that a procedure developed for isolating a protein from natural sources may not work successfully with the recombinant product, because the nature of the other proteins present influences many fractionation procedures.

Because of the difficulties often experienced in purifying recombinant proteins, a variety of vector systems (see *UNIT 6.1* and Sassenfeld, 1990, for some examples) have been developed in which the expressed product is a fusion protein containing an N-terminal polypeptide that simplifies purification. Such "tags" can be subsequently removed using a specific protease. A further advantage is that the expression level is dictated mainly by the transcription and translation signals for the fusion portion of the protein, which are optimized. Tags used include proteins and polypeptides for which there is a specific antibody, binding proteins that will interact with columns containing a specific ligand, polyhistidine tags with affinity to immobilized metal columns, sequences that result in biotinylation by the host and enable purification on an avidin column, and sequences that confer insolubility under specified conditions (*UNIT 6.1 & UNIT 6.5*).

Unstable proteins may be modified by the molecular biological technique of site-directed mutagenesis to remove the site of instability—for instance, an oxidizable cysteine. Such techniques are appropriate for commercial production of proteins, but may of course alter natural functioning parameters. Increased thermostability can be one modification, although it is not easy to predict mutations that will improve that parameter. Thermostable proteins originating from thermophilic bacteria do not need structural modification and, if expressed in large amounts, can be purified satisfactorily in one step by simply heat-treating the extract at 70°C for 30 min, which denatures virtually all the host proteins (e.g., see Oka et al., 1989).

The host bacteria used for production of

DO NOT COPY
UNIT 1.2
LIBRARY

recombinant proteins are usually *Escherichia coli*, or *Bacillus subtilis*; they may express proteins at 1% to over 50% of the cellular protein, depending on such variables as the source, promoter structure, and vector type. Generally the proteins are expressed intracellularly, but leader sequences for excretion may be included. In the latter case, the protein is generally excreted into the periplasmic space, which limits the amount that can be produced. Excretion from gram-positive species such as *B. subtilis* sends the product into the culture medium, with little feedback limitation on total expression level.

Literature Cited

Oka, M., Yang, Y.S., Nagata, S., Esaki, N., Tanaka, M., and Soda, K. 1989. Overproduction of thermostable leucine dehydrogenase of Bacillus stearothermophilus and its one-step purification from recombinant cells of *Escherichia coli*. *Biotechnol. Appl. Biochem.* 11:307-316.

Sassenfeld, H.M. 1990. Engineering proteins for purification. *Trends Biotechnol.* 8:88-93.

Key References

Deutscher, M.P. (ed.) 1990. Guide to protein purification. *Methods Enzymol.* 182:1-894.

Extensive collection of purification methods with some general protocols and examples.

Janson, J.-C. and Ryden, L.G. 1989. Protein Purification: Principles, High Resolution Methods, and Applications. VCH Publishers, New York.

A useful collection of methods and examples.

Kennedy, J.F. and Cabral, J.M. (eds.) 1993. Recovery Processes for Biological Materials. John Wiley & Sons, New York.

A useful introduction to the problems of large-scale methods.

Kenny, A. and Fowell, S. (eds.) 1992. Practical protein chromatography. *Methods Mol. Biol.* 11:1-327.

Extensive descriptions of affinity chromatographic techniques with protocols and recipes.

Scopes, R.K. 1993. Protein Purification, Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.

General principles of all the main techniques used in purifying proteins. A useful laboratory handbook; does not include recipes or procedures for specific proteins.

Contributed by R. K. Scopes
La Trobe University
Bundoora, Australia

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Protein Purification Flow Charts

Protein purification flow charts are presented to give a broad outline of the methods used for different types of proteins. They cannot give any detail, as the process appropriate for each protein will have its own variations at each stage. In most cases, the first stage is to obtain a solution containing the desired protein, after which it can be dealt with by the many separation techniques described in the following chapters. In some cases the insolubility of the desired protein can be exploited by removing soluble fractions. Purification procedures are commonly divided into three stages: (a) the primary steps, which deal with crude mixtures of proteins and other molecules present in the raw material; (b) the secondary processing, which generates a product near to homogeneity; and (c) the polishing steps, which remove minor contaminants, a process that is especially important for therapeutic proteins.

SOLUBLE RECOMBINANT PROTEINS

Proteins expressed in a recombinant manner may be (1) soluble in the cytoplasm, (2) insoluble as inclusion bodies (see section on Insoluble Recombinant Proteins), (3) excreted from the cells into the culture medium, (4) excreted into the periplasmic space (e.g., in gram-negative bacteria), or (5) associated with organelles or membrane fractions. In addition they may be expressed (6) as the normal, mature, naturally occurring protein, (7) containing a natural leader peptide that would normally be processed, (8) as a fusion protein with a peptide that is not natural to the protein, or (9) lacking glycosylation or other post-translational modification, or incorrectly modified. Possibilities (1) to (5) affect the method of extraction used to obtain the starting material for purification. Cases (6) to (9) can affect the methods used for purification.

The scheme for purifying soluble recombinant proteins is outlined in Figure 1.3.1. The first stage is to obtain a clarified solution containing the desired protein, with as little in the way of unwanted proteins as possible. For soluble cytoplasmic proteins, case (1), it is not normally possible to exclude any significant amount of unwanted soluble proteins, but in cases (2) to (5) the compartmentalization away from the cytoplasm allows such separation in the initial stage.

It may be necessary to carry out a concentration step before proceeding, especially if the protein has been excreted into the culture medium. Normally ultrafiltration is used, although other techniques are possible, especially if the extract contains particulates that block ultrafiltration membranes.

Recombinant expression in the cytoplasm of bacteria, followed by extraction via total cell disruption, results in large amounts of nucleic acids being solubilized with the protein. A number of treatments to remove nucleic acids are possible. Streptomycin is used to precipitate ribosomal material, and cationic polymers such as protamine (a basic protein) and polyethylenimine will form insoluble complexes (at low ionic strength) with nucleic acids. In addition, viscosity caused by DNA can be reduced by adding small amounts of DNase.

INSOLUBLE RECOMBINANT PROTEINS

It has been found that many proteins expressed in bacteria (mainly in *Escherichia coli*) do not fold correctly, and as a result aggregation occurs, leading to large insoluble inclusion bodies within the cytoplasm of the cells (see Chapter 6). Although this creates major difficulties in obtaining satisfactory amounts of active native product, it greatly simplifies the initial stage of purification.

The purification scheme for recombinant insoluble proteins is outlined in Figure 1.3.2. After cell disruption, inclusion bodies can be obtained in a fairly pure state by differential centrifugation. They must now be solubilized, however, and the active protein generated by encouraging correct folding. Solubilization is usually accomplished with guanidine hydrochloride and/or urea, and thiols such as 2-mercaptoethanol or glutathione are included to disrupt any disulfides that have formed and prevent more from forming. Folding the protein correctly may require a variety of additions to the solution, as well as slow removal of the denaturant. The latter can be carried out by simple dilution or by dialysis. Folding occurs best at low protein concentrations, so dilution may be adequate. If the native protein does contain disulfides, then it is important to create redox conditions such that some (but not excessive) oxidation of thiols can occur. A combination of oxidized and reduced glutathione is commonly used. In addition, the action of the enzyme protein disulfide isomerase, which can

Copyright © 1995 by John Wiley & Sons, Inc.

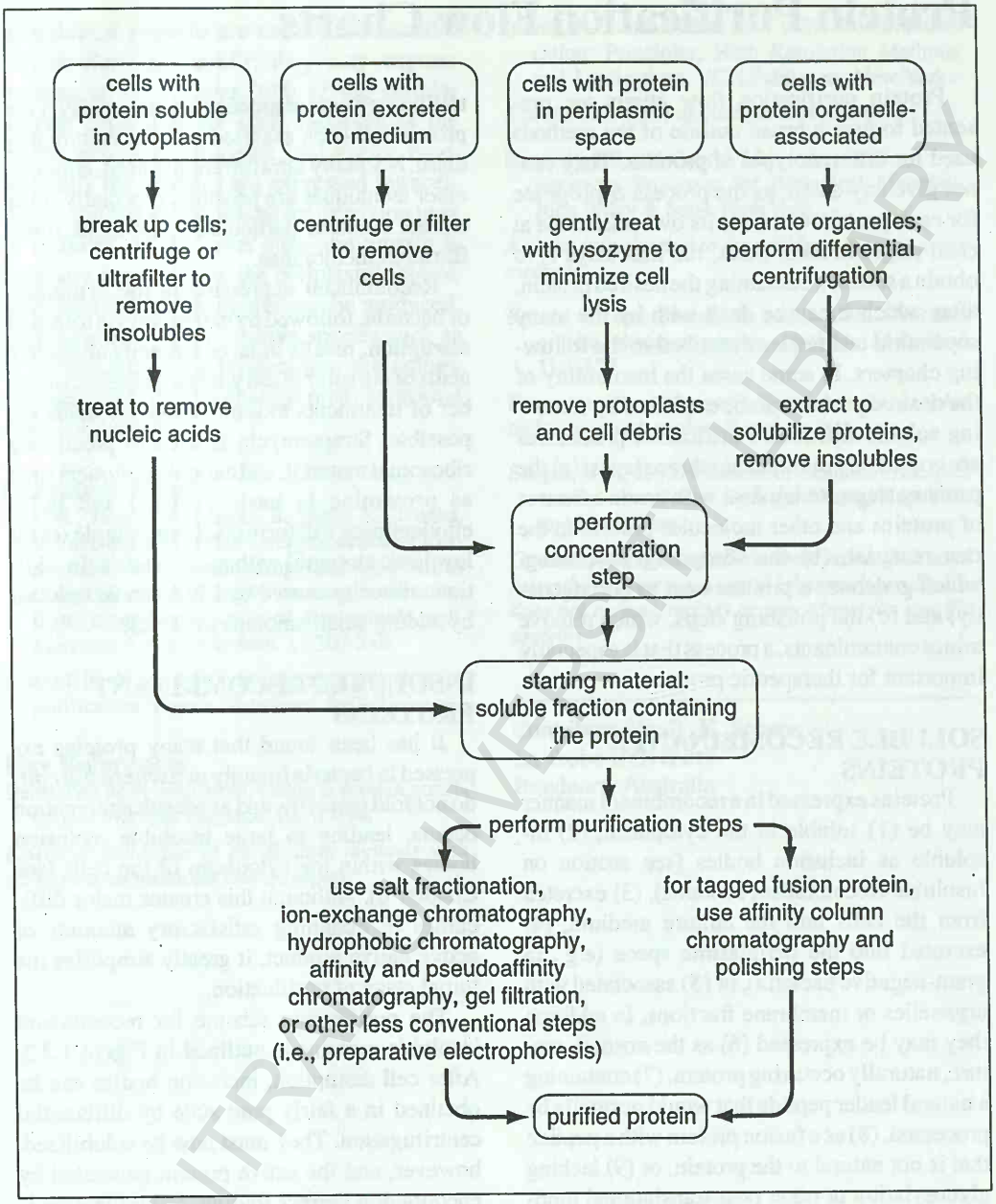

cells with protein soluble in cytoplasm
cells with protein excreted to medium
cells with protein in periplasmic space
cells with protein organelle-associated
break up cells; centrifuge or ultrafilter to remove insolubles
centrifuge or filter to remove cells
gently treat with lysozyme to minimize cell lysis
separate organelles; perform differential centrifugation
treat to remove nucleic acids
remove protoplasts and cell debris
extract to solubilize proteins, remove insolubles
perform concentration step
starting material: soluble fraction containing the protein
perform purification steps
use salt fractionation, ion-exchange chromatography, hydrophobic chromatography, affinity and pseudoaffinity chromatography, gel filtration, or other less conventional steps (i.e., preparative electrophoresis)
for tagged fusion protein, use affinity column chromatography and polishing steps
purified protein

Figure 1.3.1 Purification scheme for soluble recombinant proteins, which may be excreted or located in the periplasm, in the membrane fraction, or most commonly the cytoplasm. The first step is to obtain an extract containing the desired protein in soluble form. After this, conventional purification steps may be carried out, or affinity purification of tagged fused proteins can be performed.

make and unmake disulfides by exchange reactions, has been found to be beneficial in many cases. If the native protein is of intracellular origin, it probably will not contain disulfides; it will, however, contain cysteines, so a full reducing potential should be maintained. Specific methodology is discussed in *UNIT 6.5*.

Not all proteins can fold unassisted by other cellular components. Chaperonins are proteins whose role is to refold denatured proteins such as those that form during heat shock (Zeilstra-Ryalls et al., 1991). The most studied, and just becoming commercially available as of 1995, are the *E. coli* chaperonins GroEL and GroES, both of which are needed, together with ATP, to renature many proteins. Proline residues can adopt two isomeric conformations in proteins, and the wrong conformation is switched to the correct one by the enzyme prolyl isomerase, aiding the process of protein folding. At present these are not large-scale prospects, both because of the cost of the chap-

RCU IASRI / CENTRAL UNIVERSITY LIBRARY

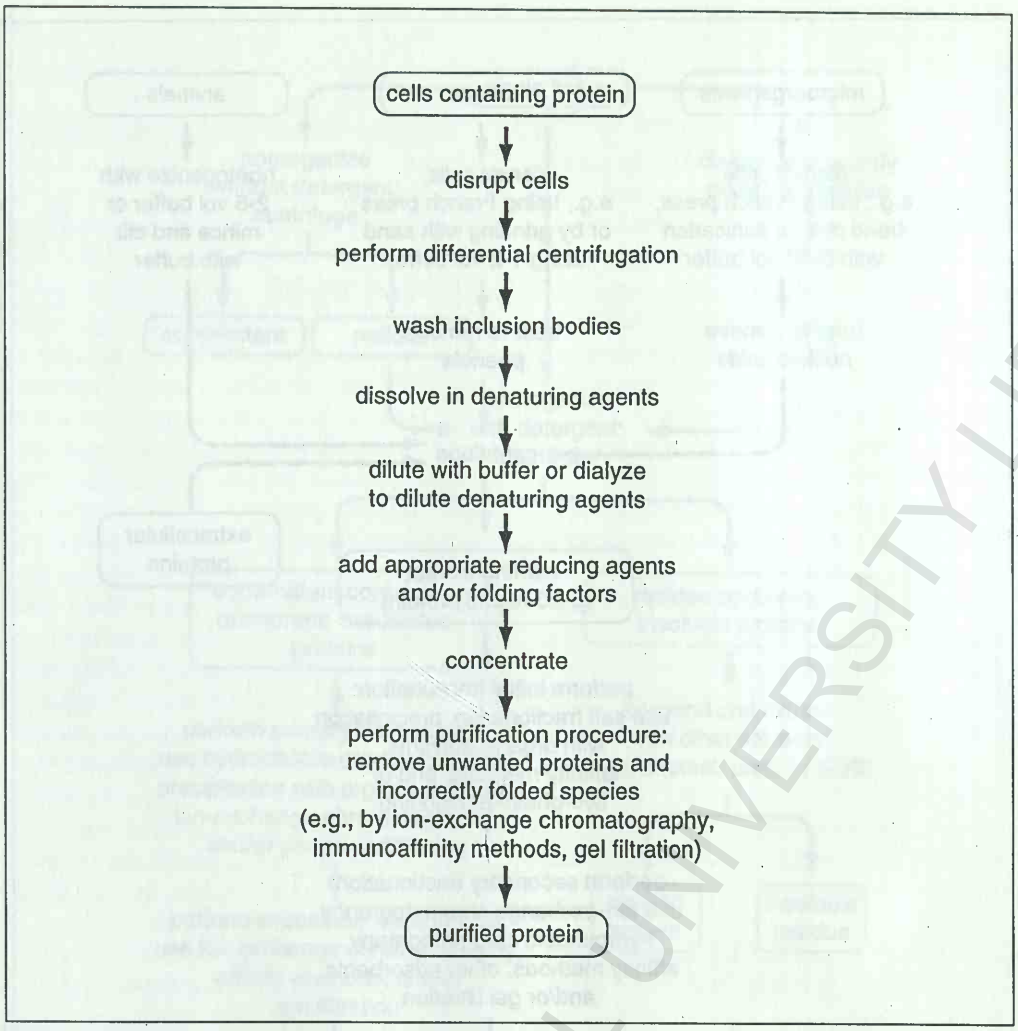


Figure 1.3.2 Purification scheme for insoluble recombinant proteins that are produced as inclusion bodies in the cytoplasm of host cells. The cells must be broken open, and then the insoluble inclusion bodies are separated by differential centrifugation. Solubilization is achieved by the use of denaturing solvents, and renaturation of the dissolved protein occurs on removal of the denaturant. Further polishing steps will be needed to remove small amounts of contaminating proteins as well as incorrectly folded species. Additional information can be found in *UNIT 6.3-6.5.*

eronins and because the agents operate best in vitro at very low protein concentrations.

Once the proteins are folded, the purification process consists of removing small amounts of still incorrectly folded protein plus any other host proteins that were trapped with the original inclusion bodies. The former may be difficult, since incorrectly folded species have a size and charge similar to those of the correct product. However, subtle differences arising from the folded conformation can be exploited by chromatographic techniques. In ideal cases immunoaffinity techniques using antibodies specific for either the incorrectly folded form or the correct one can be used to resolve the mixture.

SOLUBLE NONRECOMBINANT PROTEINS

There are so many sources of soluble proteins that it is not possible to give a complete overview of methods used to obtain starting extracts from which a desired protein can be isolated. The sources can be classified as either microorganisms, plants, or animals, as shown in Figure 1.3.3, but these in turn should be subdivided according to how the starting extract is obtained. In particular there is a distinction between extracellular and intracellular proteins. With the latter it is necessary to disrupt the cells and release the proteins, whereas with the former, if the extracellular fluid can be obtained directly, there need be no contamination with intracellular proteins. Extracellular

BOSTON UNIVERSITY LIBRARY

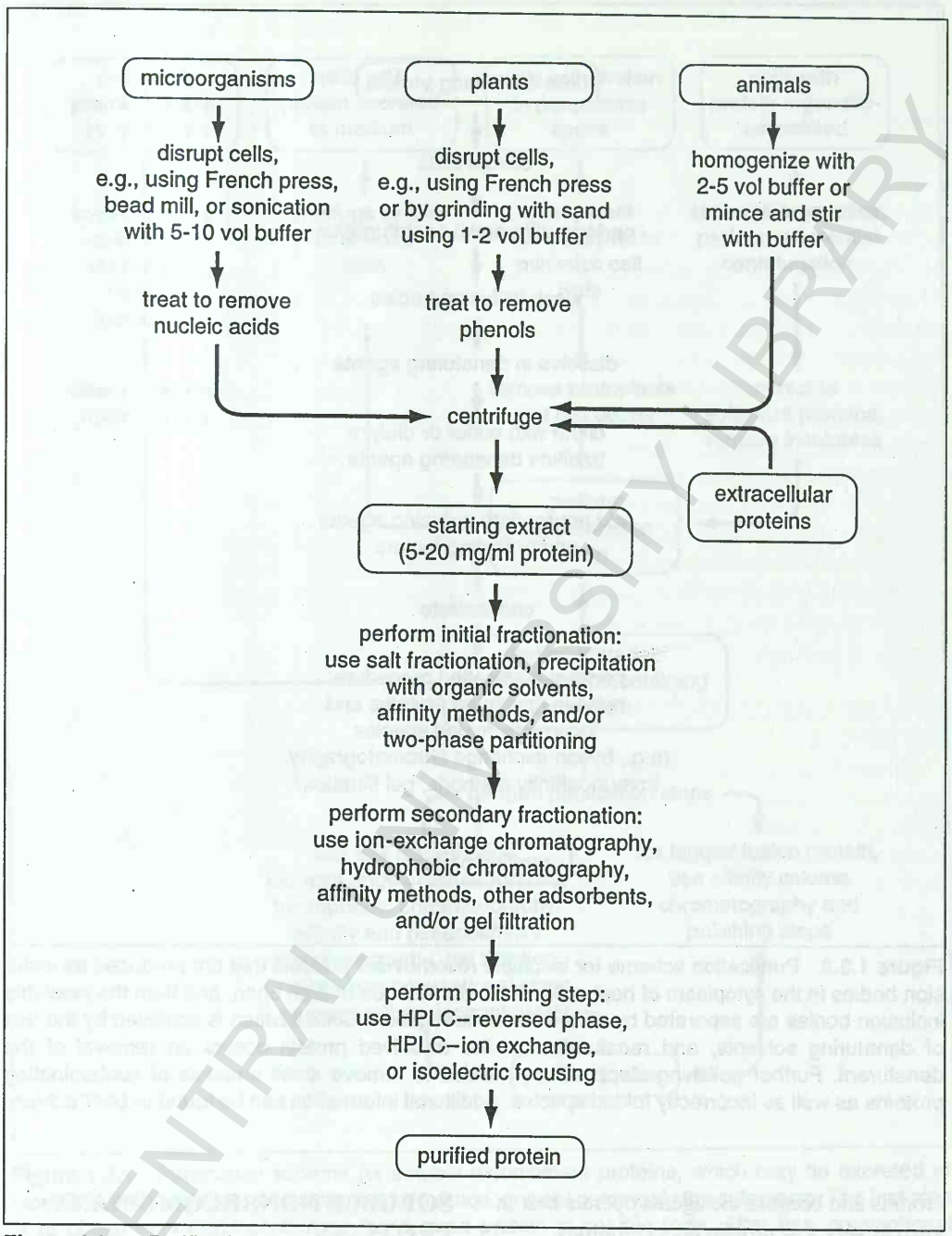

microorganisms
plants
animals
disrupt cells, e.g., using French press, bead mill, or sonication with 5-10 vol buffer
disrupt cells, e.g., using French press or by grinding with sand using 1-2 vol buffer
homogenize with 2-5 vol buffer or mince and stir with buffer
treat to remove nucleic acids
treat to remove phenols
centrifuge
extracellular proteins
starting extract (5-20 mg/ml protein)
perform initial fractionation: use salt fractionation, precipitation with organic solvents, affinity methods, and/or two-phase partitioning
perform secondary fractionation: use ion-exchange chromatography, hydrophobic chromatography, affinity methods, other adsorbents, and/or gel filtration
perform polishing step: use HPLC–reversed phase, HPLC–ion exchange, or isoelectric focusing
purified protein

Figure 1.3.3 Purification scheme for soluble proteins present in their natural host cells. Cells must be disrupted to release the proteins, usually in the presence of 2 to 10 ml of a suitable buffer per gram weight. After removal of insoluble material, the process will generally require several steps, using various standard fractionation procedures in a suitable order. For production of highly pure protein, a final polishing step may be required to remove final trace contaminants. Additional information can be found in *UNIT 6.2*.

sources include microorganism culture medium, plant and animal tissue culture medium, venoms, milk, blood, and cerebrospinal fluids. Soluble proteins may also occur within organelles such as mitochondria; these may be best obtained by first isolating the organelle, then disrupting it to release the contents.

The starting extract normally contains between 5 and 20 mg protein per milliliter, though lesser concentrations can be dealt with, especially if working on a small scale. It may be necessary to include a concentration step before starting the purification process in order to approach that level. There are exceptions to every rule, however, and very high protein concentrations can be handled, for example,

RCU IAS CENTRAL UNIVERSITY LIBRARY

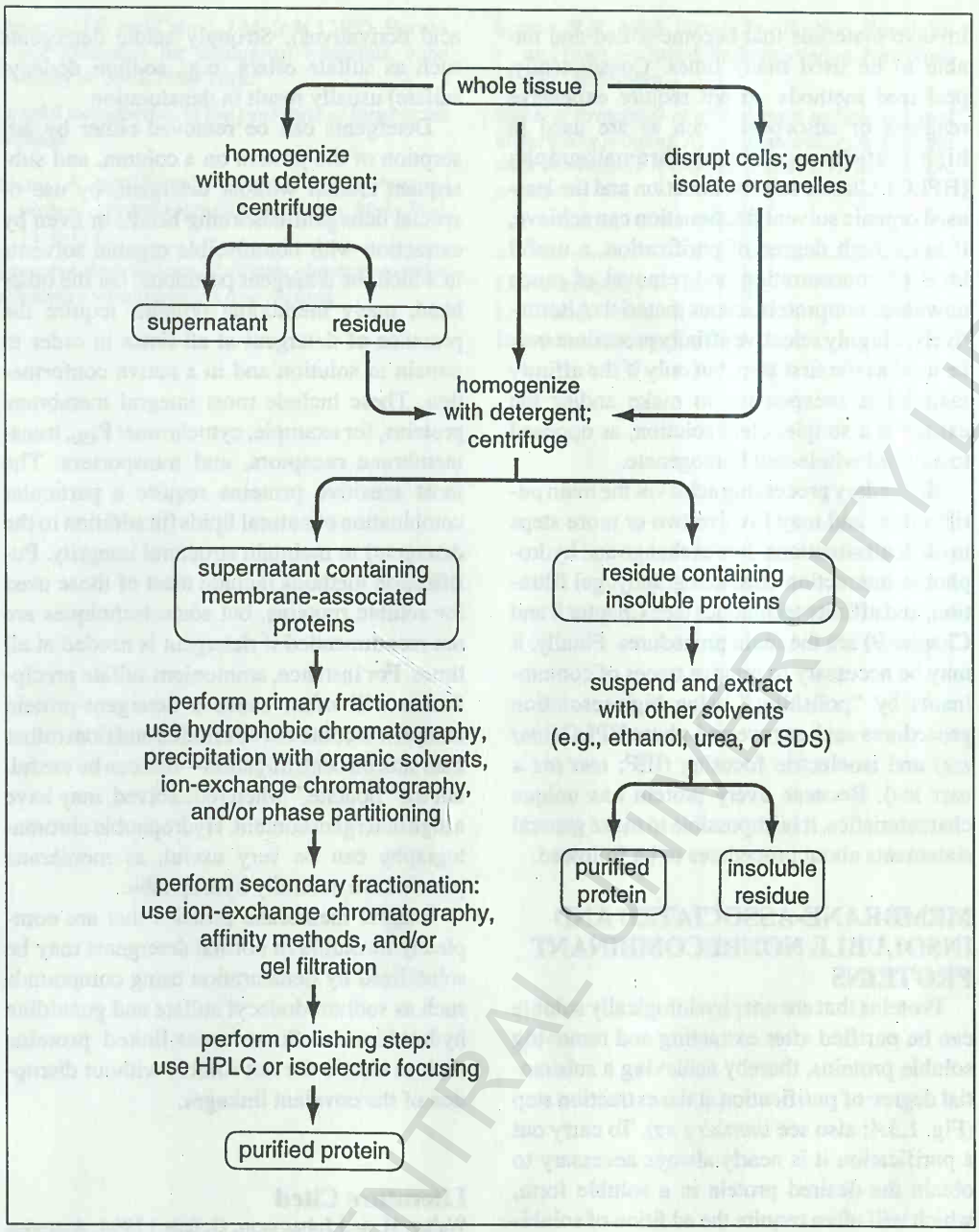

whole tissue
homogenize without detergent; centrifuge
disrupt cells; gently isolate organelles
supernatant
residue
homogenize with detergent; centrifuge
supernatant containing membrane-associated proteins
residue containing insoluble proteins
perform primary fractionation: use hydrophobic chromatography, precipitation with organic solvents, ion-exchange chromatography, and/or phase partitioning
suspend and extract with other solvents (e.g., ethanol, urea, or SDS)
perform secondary fractionation: use ion-exchange chromatography, affinity methods, and/or gel filtration
purified protein
insoluble residue
perform polishing step: use HPLC or isoelectric focusing
purified protein

Figure 1.3.4 Purification scheme for membrane-associated and poorly soluble proteins (nonrecombinant). An initial purification can be achieved by isolation of organelles containing the desired protein. Membrane proteins are normally solubilized with a nonionic detergent, although loosely associated proteins may be extracted without detergent at high pH, with EDTA, or with small amounts of an organic solvent such as N-butanol. Normal fractionation procedures may need some modification if the detergent is required throughout to maintain the integrity of the protein.

with two-phase partitioning (Walter and Johansson, 1994). When isolating proteins on a large scale, the volumes being manipulated become of increasing concern, so maximizing protein concentration can be an important aim. The starting extract should be clarified, usually by centrifugation; on a large scale, ultrafiltration methods are becoming more widely used. Pretreatment of certain extracts to remove excessive amounts of nucleic acids, phenolics, and lipids may be necessary in order to obtain an extract that is amenable to standard fractionation procedures.

Fractionation procedures can somewhat arbitrarily be divided into three steps: initial fractionation, secondary fractionation, and polishing. Initial processing, which deals with a large amount of extract that is not all protein, may

BOSTON CENTRAL UNIVERSITY LIBRARY

involve materials that become soiled and unable to be used many times. Consequently, preferred methods do not require expensive reagents or adsorbents such as are used in high-performance liquid chromatography (HPLC). Classic salt fractionation and the less-used organic solvent fractionation can achieve, if not a high degree of purification, a useful level of concentration and removal of much unwanted nonproteinaceous material. Alternatively, a highly selective affinity procedure may be used as the first step, but only if the affinity material is inexpensive to make and/or the extract is a simple, clear solution, as opposed to a turbid whole-cell homogenate.

Secondary processing achieves the main purification, and may involve two or more steps in difficult situations. Ion-exchange and hydrophobic-interaction chromatography, gel filtration, and affinity techniques (see Chapter 8 and Chapter 9) are the main procedures. Finally, it may be necessary to remove traces of contaminants by "polishing," using high-resolution procedures such as reversed-phase HPLC (*UNIT 11.6*) and isoelectric focusing (IEF; *UNIT 10.2 & UNIT 10.4*). Because every protein has unique characteristics, it is impossible to make general statements about procedures to be followed.

MEMBRANE-ASSOCIATED AND INSOLUBLE NONRECOMBINANT PROTEINS

Proteins that are not physiologically soluble can be purified after extracting and removing soluble proteins, thereby achieving a substantial degree of purification at the extraction step (Fig. 1.3.4; also see *UNITS 6.1 & 6.2*). To carry out a purification it is nearly always necessary to obtain the desired protein in a soluble form, which will often require the addition of solubilizing agents such as detergents. Some proteins remain insoluble even with detergent treatment, and so can be substantially purified by removing the soluble fractions. Some membrane-associated proteins become partly solubilized during breaking up of the tissue, and recovery in the particulate fraction may be poor. In such cases it may be best to solubilize the whole tissue by including detergent in the homogenizing buffer. Extraction of insoluble residues using detergents can be done differentially; some proteins are released at low detergent concentration, whereas others require complete solubilization of the membrane fraction. Suitable detergents include nonionic (e.g., Triton) and weakly acidic types (e.g., cholic acid derivatives). Strongly acidic detergents such as sulfate esters (e.g., sodium dodecyl sulfate) usually result in denaturation.

Detergents can be removed either by adsorption of the protein on a column, and subsequent elution without detergent, by use of special detergent-adsorbing beads, or even by extraction with nonmiscible organic solvents in which the detergent partitions. On the other hand, many membrane proteins require the presence of detergent at all times in order to remain in solution and in a native conformation. These include most integral membrane proteins, for example, cytochrome: P_{450}, transmembrane receptors, and transporters. The most sensitive proteins require a particular combination of natural lipids (in addition to the detergent) to maintain structural integrity. Purification methods include most of those used for soluble proteins, but some techniques are not recommended if detergent is needed at all times. For instance, ammonium sulfate precipitation will often cause a detergent-protein complex to come out of solution and float rather than sink on centrifugation—this can be useful, but the "floatate," when redissolved, may have a high detergent content. Hydrophobic chromatography can be very useful, as membrane proteins are naturally hydrophobic.

Integral membrane proteins that are completely insoluble in normal detergents may be solubilized by denaturation using compounds such as sodium dodecyl sulfate and guanidine hydrochloride. Some cross-linked proteins such as elastin are not soluble without disruption of the covalent linkages.

Literature Cited

Walter, H. and Johansson, G. (eds.) 1994. Aqueous two-phase systems. *Methods Enzymol.* 228:1-725.

Zeilstra-Ryalls, J., Fayet, O., and Georgopoulos, C. 1991. The universally-conserved GroE (Hsp60) chaperonins. *Annu. Rev. Microbiol.* 45:301-325.

Key References

Deutscher, M.P. (ed.) 1990. Guide to protein purification. *Methods Enzymol.* 182:1-894.

Extensive collection of purification methods, with some general protocols and examples.

Janson, J.-C. and Ryden, L.G. 1989. Protein Purification: Principles, High Resolution Methods, and Applications. VCH Publishers, New York.

A useful collection of methods and examples.

RCU IAS LIBRARY

Kennedy, J.F. and Cabral, J.M. (eds.) 1993. Recovery Processes for Biological Materials. John Wiley & Sons, New York.

A useful introduction to the problems of large-scale methods.

Kenny, A. and Fowell, S. (eds.) 1992. Practical protein chromatography. *Methods Mol. Biol.* 11:1-327.

Extensive descriptions of affinity chromatographic techniques with protocols and recipes.

Scopes, R.K. 1993. Protein Purification, Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.

General principles of all the main techniques used in purifying proteins. A useful laboratory handbook; does not include recipes or procedures for specific proteins.

Contributed by R. K. Scopes
La Trobe University
Bundoora, Australia

BCU IASI / CENTRAL UNIVERSITY LIBRARY

COMPUTATIONAL ANALYSIS

2

INTRODUCTION		2.0.1
2.1	**Computational Methods for Protein Sequence Analysis**	2.1.1
	Theoretical Background for Protein Sequence Analysis	2.1.1
	Matrix Methods for Sequence Comparison: Dot Plots	2.1.3
	Sequence Similarity Searching	2.1.3
	Alignment Method Considerations	2.1.6
	Scoring Matrices	2.1.7
	Multiple Alignments	2.1.8
	Cluster Methods and Trees	2.1.10
	Identification of Functional Sites	2.1.12
	Internet Resources	2.1.14
	Literature Cited	2.1.14
2.2	**Hydrophobicity Profiles for Protein Sequence Analysis**	2.2.1
	Introduction	2.2.1
	Methodology	2.2.2
	Applications	2.2.5
	Conclusions	2.2.10
	Accessibility of Software	2.2.11
	Literature Cited	2.2.11
2.3	**Protein Secondary Structure Prediction**	2.3.1
	Overview of Prediction Schemes	2.3.1
	Methods for Secondary Structure Prediction	2.3.3
	Application of Secondary Structure Prediction Methods	2.3.8
	Comparison of Prediction Schemes	2.3.9
	Recommendations and the Future	2.3.16
	Conclusions	2.3.16
	Accessibility of Software	2.3.17
	Literature Cited	2.3.18
2.4	**Internet Basics**	2.4.1
	Connecting to the Internet	2.4.1
	Electronic Mail	2.4.3
	File Transfer Protocol	2.4.8
	Gopher	2.4.12
	The World Wide Web	2.4.13
	Note	2.4.18
	Literature Cited	2.4.18

continued

BCU IAST / CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

2.5 Sequence Similarity Searching Using the BLAST Family of Programs 2.5.1
Obtaining Blast Documentation and Help 2.5.1
Submitting Sequences for Database Homology Searching 2.5.2
Basic Protocol 1: Submitting a Sequence to BLAST Using the Electronic Mail Server 2.5.3
Basic Protocol 2: Submitting a Sequence to BLAST Using the Command-Line Client Program 2.5.4
Basic Protocol 3: Submitting a Sequence Using NetBLAST Graphical Client Program 2.5.5
Basic Protocol 4: Submitting A Sequence Using the WWW Interface 2.5.5
Example 1: Using BLASTN to Search a Nucleotide Sequence Database with a DNA Sequence 2.5.6
Example 2: Using BLASTX to Search a Protein Sequence Database with Open Reading Frames Derived from Conceptual Translations of a DNA Query Sequence 2.5.10
Example 3: Using BLASTP to Search a Protein Sequence Database with a Protein Sequence Query Using the Filter Option to Mask Low-Complexity Subsequences 2.5.14
Example 4: Using the WWW Interface 2.5.19
Example 5: Using TBLASTN to Search a Translated Nucleotide Database with a Protein Sequence 2.5.20
Example 6: Using the NetBLAST Graphical Client Program 2.5.25
Example 7: Using TBLASTX to Search Open Reading Frames Derived from Conceptual Translations of a Nucleotide Sequence Database Using Open Reading Frames Derived from a Nucleotide Sequence Query 2.5.26
Commentary 2.5.29
Strategies and Tips 2.5.34
Literature Cited 2.5.36

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 2
Computational Analysis

INTRODUCTION

The development of the techniques of modern molecular biology, including PCR, cloning, and nucleotide sequencing, has revolutionized the practice of protein chemistry. Whereas the initial phase of detailed study of a new protein activity once involved chemical analysis of the amino acid sequence by Edman degradation of multiple fragments, it is more common today to establish a protein's linear sequence from the nucleotide sequence of the gene that encodes it. While the number of new sequences reported based solely on classical protein sequence analysis has drastically declined, the molecular biology revolution has generated a rapidly expanding collection of new sequence data. The current efforts to sequence specific genomes, including the human genome, in their entirety will undoubtedly lead to further expansion of this information. Because linear sequence information is so readily available, computer methods for analyzing these data have also developed rapidly. This chapter will present methods for deriving important information on possible familial, structural, and functional relationships using only the linear sequence information as input.

As soon as any sequence information is obtained—for example, by running a newly isolated protein through an N-terminal gas-phase sequencer or by dideoxy sequencing a piece of cloned DNA—searches for similarity to other proteins can begin. If the new sequence can be identified as related to that of a known enzyme or gene from another organism, this provides excellent initial confirmation that the protein or gene under study is worthy of further investigations. Conversely, if a match is found to the sequence of a known protein that could only plausibly arise from sample contamination, the study can be halted immediately at considerable savings of time, energy, and resources. *UNIT 2.1* describes procedures for searching sequence databases, provides theoretical background to aid in interpreting the resulting output, and discusses resources for further exploration of sequence relationships.

Once substantial sequence information (particularly a whole gene sequence) is obtained, the most pressing question is what structure it is likely to have. A first step toward answering this question is the analysis of the corresponding amino acid sequence for hydrophobic and hydrophilic regions. Of particular interest are patterns or groupings of hydrophobic residues—typically transmembrane helices, leucine zippers, and amphipathic segments—which can indicate structural and even functional features of the protein. Hydropathy scales, indices designed to quantify the relative hydrophobicity of amino acid residues within a protein, and their application to protein sequences are reviewed in *UNIT 2.2*.

More sophisticated forms of analysis can be used to predict overall protein secondary structure based on linear sequence. The procedures developed for this purpose utilize statistical data that has been compiled over the past 25 years as more and more three-dimensional protein structures have become available. Although these methods are not completely accurate at present, the success rate for predictions is now in the 60% to 80% range. This information is fully adequate to develop hypotheses, design experimental tests, and suggest relationships to known tertiary structures. The prediction of protein secondary structure and folding class are reviewed in *UNIT 2.3*. As well as providing

Contributed by Ben M. Dunn
Current Protocols in Protein Science (1997) 2.0.1-2.0.2
Copyright © 1997 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

theoretical background and detailing the different predictive approaches, the discussion includes an analysis of the strengths and weaknesses of each approach with respect to different types of proteins (using human hemoglobin β, interleukin 1β, and profilin I as examples) and gives some guidance to obtaining the necessary software.

To perform any method of computer analysis, especially comparisons with other known protein sequences, it is necessary to have facile methods for transfer and capture of a variety of files. Fortunately, standard methodology now exists to perform many of these functions. *UNIT 2.4* provides detailed instructions for the initial steps of getting connected, finding information, and transferring data between computers around the world. Although each university or company will have their own local protocols and systems, and although individuals will need to obtain help from their local computer wizards, *UNIT 2.4* is written to provide directions that are generally applicable to most systems.

Specific recommendations for utilizing various forms of the BLAST programs for sequence searching and comparison are provided in *UNIT 2.5*. Protocols for submission of search commands and search sequence are provided in detail, and examples of search results are presented.

The ultimate goal of any protein chemist should be to understand the relationship between the three-dimensional structure and the functional properties. Future updates to this chapter will address methods to complete the sequence-to-structure algorithm.

Ben M. Dunn

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Computational Methods for Protein Sequence Analysis

Despite continued advances since the first applications of computers in protein sequence analysis in the late 1960s, computer-assisted biological sequence analysis is still undergoing rapid changes and improvements. With so much activity and growth in the scope and importance of genomic research, the need for efficient biological sequence analysis has become more urgent. Nevertheless, users of computational tools should be aware of the limitations of particular methods. Good results require the user to be able to select the right tools on the basis of some theoretical knowledge and to interpret all results with biological correctness in mind.

This unit is presented as a guide to addressing the issue of what to do with a protein sequence once it is obtained. The practical approach depends on how far the investigation has progressed. If only a short part of the protein (e.g., ~10 to 20 amino acids) has been determined, initial scans of databases to see if there are any significant matches with a known sequence are appropriate. If a complete protein sequence is in hand, it is time to start identifying regions of similarity that resemble functional sites of known proteins.

THEORETICAL BACKGROUND FOR PROTEIN SEQUENCE ANALYSIS

Computer representations of protein sequences are of two basic types: artificial and natural. An example of an *artificial* or symbolic protein sequence is the common 24-letter residue code sequence (ABCDEFGHIKLMNPQRSTVWXYZ*) that represents each amino acid as suggested by the Nomenclature Committee of the International Union of Biochemistry (IUB). This code allows for ambiguities and unknown residue modifications. A *natural* protein sequence, on the other hand, is directly representative of the physical properties of the individual amino acids; for example, in a natural protein sequence the individual elements might be vectors representing the physical characteristics of the amino acid (Robson and Greaney, 1992; Rohde and Bork, 1993). Most molecular sequence analysis algorithms have been designed for the IUB 24-character amino acid representation.

A sequence is an ordered list of elements. Sequences can vary in length, composition, and pattern. The number of possible different protein sequences possible of a particular length is 20^L, where L is the length. The patterns in such sequences are said to constitute the **primary structure** of the protein. In general, proteins have nonrandom sequences that are presumed to be related in the sense that all proteins have evolved from other proteins (Doolittle, 1981, 1986, 1989, 1990). Protein sequences are also subject to **sequence convergence**: different sequences may represent the same native structure of a protein.

Some terms that deal with comparisons of protein sequence data are often used and misused in the literature. **Similarity** is generally used to describe two sequences or structures that resemble one another due to the order of identical residues. **Homology** means that similar sequences or structures resemble one another due to descent from a common evolutionary ancestor. It is important to realize that the sequence analysis tools used to search data bases detect character patterns that resemble portions of the query sequence. This means that they are finding similarities. Homologies are identified through the recognition of correlated structural and functional data in addition to similarity data.

There are three basic types of sequence analyses: listings, traces, and alignments (Kruskal, 1983; Kruskal and Sankoff, 1983). Let $\mathbf{a} = (a_1,..., a_m)$ and $\mathbf{b} = (b_1,..., b_n)$ be two sequences and let E be a set of elementary operations including insertion, deletion, and substitution. An **indel** is either an insertion or a deletion. A **listing** is a finite sequence of operations taken from E that transform **a** into **b**. A **trace** is a one-to-one correspondence of a subsequence of **a** and a subsequence of **b**. An **alignment** of **a** and **b** is a one-to-one correspondence such that each element of one sequence corresponds either to an element of the opposite sequence or to a null element indicating the presence of a gap. The best alignments are obtained when the lengths of **a** and **b** are similar. The correspondence between two nonnull elements is called an **identity** or a **continuation** if the two elements are identical and a **substitution** otherwise. Lines representing the correspondence between two sequences should not cross.

Copyright © 1995 by John Wiley & Sons, Inc.

In the comparison of two different alignments, it is not necessarily the one with the larger number of identities that is better. An alignment with fewer indels (<5%) might be preferable; therefore, some alignment methods set a limit to the number of indels allowed in a given alignment.

Alignments are the most frequently used type of sequence analysis. There are two basic types of alignments: local and global. **Local** alignments are used to identify the most similar common subregions of two sequences. They are especially useful in identifying distantly related sequences. A **global** alignment, on the other hand, attempts a comparison of two sequences in their entirety and, therefore, often misses small subregions of similarity (Pizzi et al., 1991). Global alignments are most useful for comparing highly similar or homologous sequences that are close to one another in length. Two sequences are said to be **homologous** if it can be inferred on the basis of all available information that the sequences have a common ancestor. There are no degrees of homology.

Divergent sequences are those that can be derived from the same sequence, usually by applying the operations in *E*. One way that divergent sequences are identified is by applying scoring matrices. There are several scoring matrices available based on allowable amino acid substitutions found in protein alignments of related proteins (see discussion of scoring matrices that follows). The most common matrices used are the protein alignment matrix (PAM; Dayhoff, 1978); the mutation data matrix (MDM; George, et al., 1990), which is a log-odds version of PAM; and BLOSUM (Henikoff and Henikoff, 1993), which incorporates the consideration of the conversation of function data between aligned proteins. It should be noted that the restriction of global alignments to operations in *E* would fail to align two sequences containing **transpositions**, in which segments of the protein sequence have been reordered within the sequence (Burks, 1990; States and Boguski, 1990).

The similarity of two biological sequences is usually specified by either a similarity measure or a distance measure. **Similarity measures** typically compute the score for an alignment using an optimizing function that rewards identities and penalizes nonidentities and gaps. **Distance measures** are usually defined as functions that satisfy the axioms for a metric function:

Distance(**a**,**b**) = 0 if and only if **a** = **b**;

Distance(**a**,**b**) = distance(**b**,**a**);

Distance(**a**,**b**) ≤ distance(**a**,**c**) + distance(**c**,**b**).

Distance functions can be thought of as minimizing differences to find the best alignments, whereas similarity methods search for the best alignments by maximizing similarities. "All problems known to be solvable with distance methods can be solved with similarity methods" (Waterman, 1989). The metric approach appeals to the geometric intuition: sequences can be thought of as points in space, and any three distinct points **a**, **b**, and **c** form a triangle *ABC* that satisfies the rules above. The smaller the distance between two points, the greater the similarity of the sequences.

There are many factors to be considered in protein sequence comparison (Hodgman, 1992). Amino acids are not equally abundant in protein sequences. Moreover, proteins fold into structures having well-conserved hydrophobic interior residues and poorly conserved exterior residues. Indeed, there is a descending order of residue conservation: active-site residues > crucial folding residues > internal packing residues > surface residues. Some residues have preference for position, secondary structure, and substitution; for example, active-site residues are often found in loops that are part of a supersecondary unit and may be partially buried. An identity of >25% implies similarity of function, whereas 18% to 25% identity may indicate similarity of structure or function. The latter range is within the "Twilight Zone" (15% to 25% identity) that, according to Doolittle, arises in the theory of protein evolution (Doolittle, 1986).

Alignment methods usually have a specific region of applicability (Taylor, 1988). For example, pairwise alignment is most applicable at high levels of identity (~70%); consensus methods (~45%), template methods (~25%), and structure prediction methods are applicable at even lower levels of identity. These ranges take into account the fact that even two unrelated or random sequences over 100 residues long may show identity in the range of 20%.

Two complementary properties of sequence comparison methods are sensitivity and specificity (States and Boguski, 1990). **Sensitivity** is the ratio of false negatives to the sum of true positives and false negatives; **specificity** is the ratio of true positives to the sum of true positives and false positives.

The goal of this unit is to provide some

starting points for the analysis of protein sequence data. A theoretical review of principles is complemented by practical examples of computer methods for protein sequence analysis that were done using the GCG package of programs (Genetics Computer Group, 1994). There are many other commercial packages for IBM-compatible computers, Macintoshes, and UNIX workstations that have similar capabilities. There are also many public domain software packages available over the Internet that have identical or enhanced functionalities (see Table 2.1.1). When starting to assemble programs for protein sequence analysis, one is strongly advised to browse the Internet before purchasing. The CD-ROM version of *Current Protocols in Protein Science* also includes the Entrez and MACAW multiple sequence alignment programs from the National Center for Biotechnology Information (NCBI) at no additional charge. *UNIT 2.2* describes the analysis of linear protein sequence with respect to hydrophobic/hydrophilic characterization that can suggest structural or functional segments. *UNIT 2.3* presents methods for the prediction of secondary structural features of linear protein sequences.

MATRIX METHODS FOR SEQUENCE COMPARISON: DOT PLOTS

Many computer-based sequence comparison methods rely on matrix methods (Fitch, 1966, 1969; Gibbs and McIntyre, 1970; McLauchlan, 1971, 1972; Maizel and Lenk, 1981; Argos, 1987; Landau et al., 1988, 1990). Pairwise comparison of two sequences can be represented in its simplest form by a matrix that is the logical product of the two sequences: if a corresponding pair of elements is identical, then the matrix element at the intersection of the row and column is marked true (dot); if the pair is different, then the matrix element is marked false (blank). The result is a visualization of the similarity of the two sequences.

In dot-plot matrix displays, a run of identities will be seen as a continuous diagonal segment within the matrix. Discontinuities along diagonals with horizontal shifts to parallel lines are interpreted as indels and indicate the need for gaps in any alignment. If the two sequences being compared are identical, the matrix will reveal any internal repeats and the main diagonal will be filled in. Self-comparisons always result in symmetrical dot plots.

Another commonly used dot-matrix method compares fixed-length segments (**windows**) of the sequences instead of individual residues. See Staden (1994a-d) for a description of how to use SIP, a program based on dot-matrix plots.

A **palindrome** is a subsequence that remains the same after being reversed. On a dot plot, palindromes are displayed as segments perpendicular to the main diagonal of the matrix. Palindromes are not usually relevant to protein sequences. In RNA and DNA sequences, a palindrome is defined as an inverted, self-complementary region such as ACTGATATCAGT.

More sophisticated comparisons have been defined on dot matrices that filter out isolated dots or small segments as noise. The **stringency** of the similarity determines the amount of data that will be filtered out before the matrix is displayed or undergoes further processing. In this case, the matrix elements are filled in with weighted numerical scores instead of simple identity values, and the result is plotted as a surface to visualize the regions of greatest similarity. Dot-matrix methods remain an active area of research (Streletc et al., 1991; Landes et al., 1993; Nedde and Ward, 1993).

Dot plots are the simplest form of sequence comparison using a matrix display. These graphical methods allow quick observation of the level and extent of similarity between two sequences. For example, a comparison of horse and rat serum albumin protein sequences from the SWISS-PROT protein database is summarized in Figure 2.1.1.

Dot-plot analysis is done in two phases: pairwise sequence comparison to generate a matrix of similarity values is followed by generation of the graphic display. The GCG package implements this process by using two programs—COMPARE and DOTPLOT—to generate the dot-matrix plot. COMPARE does a pairwise analysis to define positions of similarity between the two sequences. The parameters to be varied are (1) window size, or the number of residues to be compared, and (2) stringency, or the minimal number of matches per window to score a dot. DOTPLOT generates the graphic matrix display of the COMPARE analysis (Fig. 2.1.1).

SEQUENCE SIMILARITY SEARCHING

Once a small portion of protein sequence data from an unknown sample has been generated, the immediate impulse is to ask what other proteins in the database does the unknown resemble, and how similar are the sequences. Several available tools can help answer these

PANNING LIBRARY

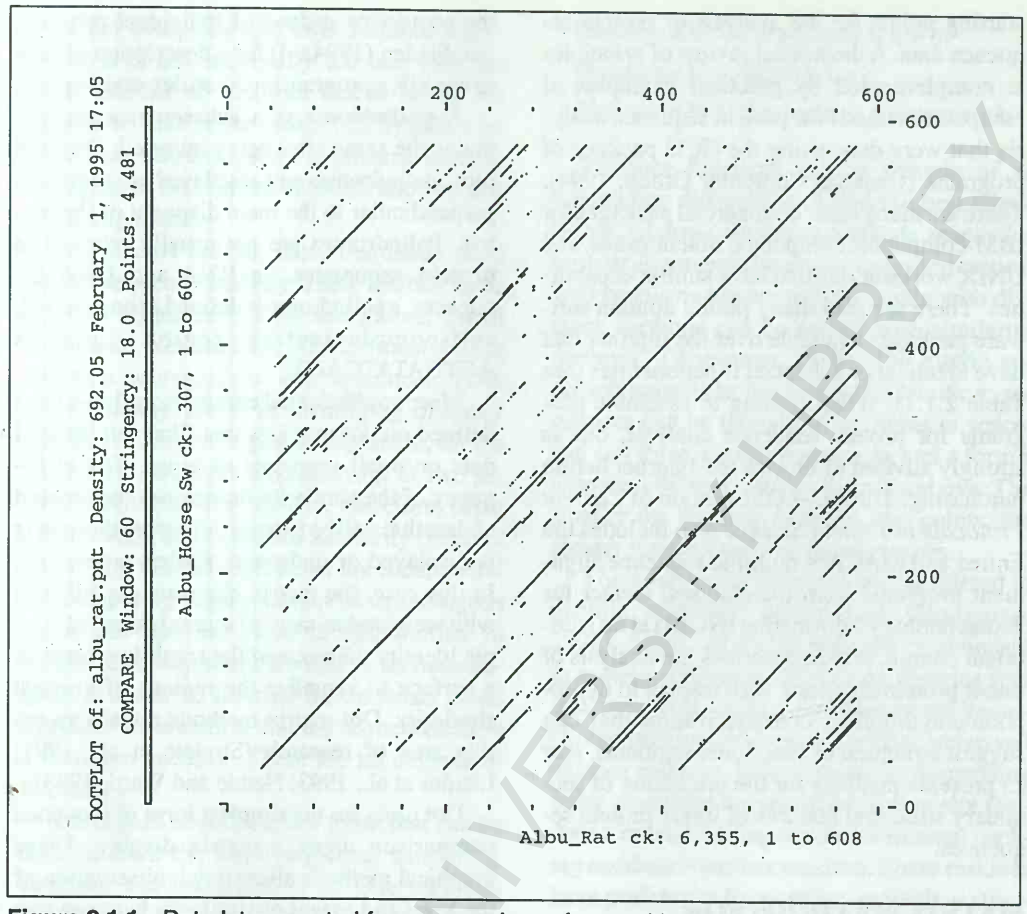


Figure 2.1.1 Dot plot generated from comparison of rat and horse albumin protein sequences. The continuous diagonal line of identity traverses the middle of the matrix. Regions of similarity between the two sequences appear as lines parallel and offset to the line of identity. Axis label takes the form filename ck: checksum_value start_position to end_position.

questions. Most commercial sequence analysis packages contain programs for performing pairwise alignments of protein sequences. The most popular tools for searching databases, FASTA (FAST All) and BLAST (the Basic Local Alignment Search Tool), are included in commercial sequence analysis suites and are also freely available to be downloaded to a local computer to be used in a stand-alone manner or as network services.

BLAST Database Searching

The program BLAST, which is available from the NCBI, conducts rapid searching of sequence databases (Altschul et al., 1990). The program searches for similarity between a query sequence and sequences in a nucleotide database (using BLASTN) or peptide database (using BLASTP). Two versions of the program are available: (1) a version that will run on a local workstation and (2) client software that connects a local computer via the Internet to a high-speed search engine located at the NCBI. The latter accesses the most recent versions of sequence databases available.

BLAST uses a word/dictionary approach to compare fixed-length "words" from the query sequence against the dictionary of words from the database searched. This program is an excellent choice for quickly screening for exact matches—for example, when the end of a new protein has been sequenced. A word of caution is, however, in order: BLAST can create enormous files, especially if the query sequence has repeated residues or is similar to a repeated motif in the databases. The BLAST output file contains (1) the sorted match statistics, probability measure of a random match and the number of matched regions; (2) the alignments generated for the query and the matched sequence region (shown in Fig. 2.1.2); and (3) a copy of the search parameters (not shown). The example in Figure 2.1.2 shows output from a BLAST search conducted using a sequence fragment from an unknown mammalian blood protein as a query against the nonredundant

BSU IASRI CENTRAL UNIVERSITY LIBRARY

Similarity Table

Sequences producing High-scoring Segment Pairs:		High Score	Smallest Sum Probability P(N)	N
sp|P35747|ALBU_HORSE	SERUM ALBUMIN PRECURSOR. >pir|S340..	101	6.8e-08	1
sp|P08835|ALBU_PIG	SERUM ALBUMIN PRECURSOR (FRAGMENT)..	98	1.8e-07	1
gp|M36787|PIGALBA_1	albumin [Sus scrofa]	98	1.8e-07	1
gp|X17055|OASERALB_4	Sheep mRNA for serum albumin. [Ovi..	96	3.4e-07	1
sp|P14639|ALBU_SHEEP	SERUM ALBUMIN PRECURSOR. >pir|S069..	96	3.5e-07	1
sp|P02769|ALBU_BOVIN	SERUM ALBUMIN PRECURSOR. >pir|A388..	95	4.8e-07	1
pir|A47391|A47391	serum albumin prepropeptide - rhes..	94	6.6e-07	1
gp|M73993|BOVALBUMIN_1	albumin [Bos taurus]	94	6.6e-07	1
gp|V01222|RNALBU_4	Messenger RNA for rat preproalbumi..	92	1.3e-06	1
sp|P02770|ALBU_RAT	SERUM ALBUMIN PRECURSOR. >pir|A938..	92	1.3e-06	1
pir|S17599|S17599	albumin - human	79	1.8e-06	1
pir|S29749|S29749	serum albumin - dog	79	4.0e-06	1
gp|A15293|A15293_1	Mature HSA [Homo sapiens]	85	1.2e-05	1
gp|A00279|A00279_1	Human serum albumin [Unknown]	85	1.2e-05	1
gp|A16010|A16010_1	synthetic mature HSA [Unknown]	85	1.2e-05	1
gp|A03758|A03758_1	serum albumin [Homo sapiens]	85	1.2e-05	1
gp|M12523|HUMALBGC_1	Human serum albumin (ALB) gene, co..	85	1.2e-05	1
gp|A09561|A09561_1	human serum albumin [Unknown]	85	1.2e-05	1
sp|P02768|ALBU_HUMAN	SERUM ALBUMIN PRECURSOR. >gp|M1252..	85	1.2e-05	1
pir|A93743|ABHUS	serum albumin precursor - human >g..	85	1.2e-05	1

Alignment

```
>sp|P35747|ALBU_HORSE SERUM ALBUMIN PRECURSOR. >pir|S34053|ABHOS serum albumin
           precursor - horse >gp|X74045|ECSERALB_1 preproalbumin [Equus
           caballus]
           Length = 607

 Score = 101 (47.0 bits), Expect = 6.8e-08, P = 6.8e-08
 Identities = 22/26 (84%), Positives = 22/26 (84%)

Query:     1 TIKSEIAHVFNDLGEPVFKGLVLVAF 26
             T KSEIAH FNDLGE  FKGLVLVAF
Sbjct:    26 THKSEIAHRFNDLGEKHFKGLVLVAF 51
```

Figure 2.1.2 BLAST output table and alignment for an unknown mammalian blood protein. The Query is the sequence of interest and the sbjct is the sequence region found in the database that matches the query. Note that in searches of the nonredundant databases at the NCBI, matches are found between the test sequence and the translation of mRNAs in GenBank. Important similarities have small probability scores, i.e., $P(N) < 10^{-3}$. This analysis suggests that the unknown is an albumin, but it is not identical to any known sequence.

protein sequence database maintained at the NCBI. This database contains a compilation of protein sequences found in all of the molecular sequence databases with duplicate sequences removed. Figure 2.1.2 presents a table of the top 20 highest similarity scoring sequences and the best alignment.

FASTA Database Searching

FASTA conducts a Lipman and Pearson (1985) search for similarity between a query sequence and any FASTA-formatted database or group of sequences collected in a database by the user. FASTA is a program developed for fast searching using an algorithm that allows for the insertion of gaps during the alignment phase. This approximates evolutionary insertions and deletions during divergence. However, the gaps are generated to maximize the number of aligned residues. FASTA does a good job of generating a global alignment for a sequence.

Figure 2.1.3 and Figure 2.1.4 present results from a FASTA search of the SWISS-PROT database with a short sequence fragment from an unknown mammalian blood protein. The FASTA output contains (1) a histogram that shows the distribution of relevant word matches found between the query sequence and the database searched; (2) the match statistics; and (3) the alignments generated for the query sequence and the matched sequence region.

BIOMEDICAL LIBRARY UNIVERSITY

```
 Score Init1   Initn
         (-)    (+)
<   2    533    533:==================================================
     4      4      4:==
     6    134    134:==================================================
     8    222    222:==================================================
    10   1964   1964:==================================================
    12   5289   5289:==================================================
    14   7609   7609:==================================================
    16   9434   9434:==================================================
    18   2834   2834:==================================================
    20   4914   4914:==================================================
    22   2686   2686:==================================================
    24   1808   1808:==================================================
    26   1241   1241:==================================================
    28    721    721:==================================================
    30    375    374:==================================================
    32    261    259:==================================================
    34    138    137:==================================================
    36     57     57:==============================
    38     27     27:==============
    40     18     18:=========
    42      4      5:==+
    44      6      7:===+
    46      0      0:
    48      0      1:+
    50      3      4:==
    52      1      1:=
    54      1      1:=
    56      2      2:=
    58      0      0:
    60      0      0:
    62      0      0:
    64      0      0:
    66      0      0:
    68      0      0:
    70      0      0:
    72      0      0:
    74      0      0:
    76      0      0:
    78      0      0:
    80      0      0:
> 80      6      6:===
Mean score calculations exclude scores greater than 72
 mean initn score:  16.3 (s.d. 5.20)
 mean init1 score:  16.3 (s.d. 5.19)
 1838 scores better than 26 saved, joining threshold: 27
```

Figure 2.1.3 FASTA histogram from a search of the SWISS-PROT database for sequences similar to a fragment from an unknown mammalian blood protein. Numbers of windows at each *init1* score are plotted. Note that there are six highly significant alignments, as well as several sequences (those with *init1* scores in the 48 to 56 range) that represent protein sequences from more distantly related species.

There are three similarity scores presented in the output from FASTA: *init1*, the highest similarity score allowing conservative substitutions; *initn*, the highest score generated by linking nonoverlapping regions; and *opt*, the score for the best alignment presented. In practice the histogram would be used to determine how many sequences have highly significant levels of similarity. The table of scores allows one to decide quickly how many and which sequence alignments should be viewed.

ALIGNMENT METHOD CONSIDERATIONS

Dynamic programming is a different matrix method that is used to find the optimal global alignment of two sequences (Needleman and Wünsch, 1970; Sellers, 1974; Pearson and Miller, 1992; Panjukov, 1993). Let **a** be a sequence of length *m* and let **b** be a sequence of length *n*. Let **D** be the working matrix. Then, given a predefined scoring function, the basic steps of the alignment algorithm (Doolittle, 1986) are as follows. Assign the scores for all exact matches to the corresponding matrix el-

BU IACS CENTRAL UNIVERSITY LIBRARY

```
Alignment Score Table

The best scores are:                                          init1 initn  opt
Sw:Albu_Horse  P35747 equus caballus (horse). serum album...   99    99   100
Sw:Albu_Rat  P02770 rattus norvegicus (rat). serum albumi...   99    99    99
Sw:Albu_Pig  P08835 sus scrofa (pig). serum albumin precu...   99    99   101
Sw:Albu_Bovin  P02769 bos taurus (bovine). serum albumin ...   98    98    99
Sw:Albu_Sheep  P14639 ovis aries (sheep). serum albumin p...   95    95    96
Sw:Albu_Human  P02768 homo sapiens (human). serum albumin...   93    93    93
Sw:Albu_Chick  P19121 gallus gallus (chicken). serum albu...   56    56    77
Sw:Albu_Mouse  P07724 mus musculus (mouse). serum albumin...   55    55    55


Best Scoring Alignment

SCORES            Init1:    99 Initn:    99 Opt:   100
            84.6% identity in 26 aa overlap

                                      10        20
Pc12_P                           TIKSEIAHVFNDLGEPVFKGLVLVAF
                                 | |||||| ||||||  |||||||||
Albu_H MKWVTFVSLLFLFSSAYSRGVLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPF
              10        20        30        40        50        60
```

Figure 2.1.4 Partial FASTA alignment table and best scoring alignment for the same search illustrated in Figure 2.1.3. The table shows the best alignment scores sorted by highest *init1* score.

ements. Transform the matrix by beginning with element D_{mn} and adding its value (less any gap penalty) to D_m^{-1}, $_n^{-1}$ and all the elements directly above it or to its left. Next, select the matrix element in column $m-2$ and above row $m-1$ that contains the largest score and repeat the summing procedure. Repeat the previous step until the entire matrix is transformed. Next, determine the path from the largest score in the matrix back toward the bottom right-hand corner. The path, with any gaps as indicated by off-diagonal moves, is then used to align **a** and **b**.

This matching algorithm is used in most global alignment programs, which vary primarily in the sophistication of the scoring functions (similarity or distance). Note that this method assumes (incorrectly) that each residue has equal importance, and it ignores sequential constraints.

Local alignments are usually found using some version of the Smith-Waterman method (1981), which applies when common regions are being sought in two sequences whose similarity is unknown or limited to a few regions, as is the case in database searches. The algorithm usually produces a single highest-scoring alignment with respect to the operations in E, although it has been successfully extended to produce the second-best and other alignments (Barton, 1990, 1993; Waterman and Eggert, 1991).

In its simplest form (Barton, 1993), the local alignment method is as follows. Given two sequences **a** and **b**, and a working (comparison) matrix **H**, the matrix elements of **H** are filled in by applying a scoring rule: the score assigned to a matrix element H_{ij} is either zero or the maximum value of any of the three scores of its neighbors and predecessors that lie directly above, diagonally to the left, or directly to the left of H_{ij}, adjusted for gaps or substitution of residues. The maximum element in **H** is found and traced back through the matrix to find the alignment.

SCORING MATRICES

Scoring matrices (George et al., 1990; States and Boguski, 1990), such as the PAM, BLOSUM (Henikoff and Henikoff, 1993), and related MDM matrices (Dayhoff, 1978), are based on the compilation of frequencies of amino acid substitution in a set of homologous sequences. The data are used as an amino acid transition probability matrix. Increasing powers of the matrix correspond to increasing divergence. Lower powers are used for sequences that seem closely related and higher powers for those that are not so obviously related. Thus, PAM250 represents a matrix with 250 accepted protein mutations per 100 amino acids. The scoring matrix method has greatly improved the quality and sensitivity of alignments; however, it does have the deficiency that it ignores structure and the corresponding sequence constraints which are significant for predicting structure.

Not all sequence comparison methods are

BOSTON UNIVERSITY LIBRARY

matrix methods. Programs such as FASTP and FASTN are representative of hash-coding methods (Dumas and Nunio, 1982; Wilbur and Lipman, 1983; Lipman and Pearson, 1985; Pearson, 1990, 1994). The methods are based on fast table lookups of subsequences of a target sequence using **hash functions**. Hash functions transform strings in numbers that serve as pointers into a table containing further information. Thus, a fixed-length subsequence (or *k*-tuple) of one sequence can quickly be checked for occurrence in another sequence after a hash table of *k*-tuples and their positions in the sequence is constructed for each sequence. FASTP focuses on regions having the highest density of identities and uses the PAM250 matrix.

Readers are referred to Pearson (1994) for an introduction to using FASTA for comparing both DNA and protein sequences. The RDF2 and RSS programs for testing the statistical significance of FASTA alignments are also discussed.

Other methods for sequence comparison include **finite state machines** that are capable of recognizing regular expressions, and various statistical approaches (States and Boguski, 1990; Tyler et al., 1991; Allison et al., 1992; Brendel et al., 1992). Regular expressions are formulas such as (**a** + **b***)**c**, which represents the set of all strings that begin with the letter **a** *or* zero or more **b**'s *and* followed by a single **c**. Complicated patterns can be constructed and detected in sequences using the finite state method. Statistical approaches include the database-searching program BLAST (see above section on BLAST database searching), which approximates gapless alignments that optimize local similarity on the basis of maximal segment pairs. The BLAST strategy is to search for only those segment pairs that contain a word pair whose comparison score exceeds some threshold. It uses the PAM120 scoring matrix to allow for conservative substitutions of residues in a comparison.

Statistical methods for sequence alignment (Karlin et al., 1988, 1989, 1991; Brendel, 1992; Pevzner, 1992; Mrazek and Kypr, 1993; Staden, 1994a-d) consider the statistical significance of a variety of protein sequence properties. Such properties include the relative composition of a sequence with respect to residue types; anomalous clusters, runs, and periodic variations of charge; internal repeats; local periodicities that might indicate regularity of structure; and the relative spacings of certain residue types. For a description of routines for plotting hydrophobicity, charge, and hydrophobic moments using the program PIP, see Staden (1994a-d).

In addition to the methods described above, interest in the development of theory and tools for linguistic (formal grammar) analysis of biological sequences is growing (Collado-Vides, 1991; Michaels et al., 1993; Searls, 1993).

The next level of sequence analysis is the simultaneous comparison of more than two sequences, as described in the following section on Multiple Alignments.

MULTIPLE ALIGNMENTS

Once a group of potentially homologous sequences has been identified, the next step could be to do a **multiple alignment** that compares the new sequence with the group of sequences determined to have highly significant similarities. There are several programs that will do multiple alignments. PILEUP from the GCG package does a good job with closely related sequences in an automated fashion. MACAW from NCBI, LINEUP from GCG, and DSCE (De Rijk and De Wachter, 1993) allow the investigator to generate an alignment manually. (The CD-ROM version of *Current Protocols in Protein Science* includes MACAW and Entrez from NCBI.)

Figure 2.1.5 demonstrates a multiple alignment generated using PILEUP from five high-scoring albumin sequences in the SWISS-PROT database identified from the FASTA analysis (Fig. 2.1.4). PILEUP performs a cluster analysis based on pairwise comparisons of each sequence to be aligned. As well as the multiple alignment (Fig. 2.1.5), PILEUP produces a dendrogram (not shown) that displays the clustered relationships between the sequences. The dendrogram is not a phylogenetic tree, but a graphical representation of the clustered similarity statistics.

After the multiple alignment is generated, one can begin to infer relationships between conserved regions and function or structure. The program Pro_Anal (Eroshkin et al., 1993) and the program for fast protein block searches (Fuchs, 1994) can help in this process. Programs like OBSTRUCT (Heringa et al., 1992) are useful in identifying clusters of protein sequences on the basis of structural and sequence similarity criteria. Kanaoka et al. (1989) have described a process of aligning protein sequences on the basis of hydrophobicity. Livingstone and Barton (1993) describe a strategy to generate multiple alignments for hierarchical analysis of residue conservation.

```
                    1                                                  50
Albu_Bovin.Sw  MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EQFKGLVLIA
Albu_Chick.Sw  MKWVTLISFI FLFSSATSRN LQRFARDAEH KSEIAHRYND LKEETFKAVA
Albu_Horse.Sw  MKWVTFVSLL FLFSSAYSRG VLRRDTHKSE IAHRFNDLGE KHFKGLVLVA
Albu_Human.Sw  MKWVTFISLL FLFSSAYSRG VFRRDAHKSE VAHRFKDLGE ENFKALVLIA
Albu_Mouse.Sw  EADKSEIADR YNDLGEIHFK CAIPNLRENY GELADCCTKQ EPERNECFLQ

                    51                                                 100
Albu_Bovin.Sw  FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
Albu_Chick.Sw  MITFAQYLQR CSYEGLSKLV KDVVDLAQKC VANEDAPECS KPLPSIILDE
Albu_Horse.Sw  FSQYLQQCPF EDHVKLVNEV TEFAKKCAAD ESAENCDKSL HTLFGDKLCT
Albu_Human.Sw  FAQYLQQCPF EDHVKLVNEV TEFAKTCVAD ESAENCDKSL HTLFGDKLCT
Albu_Mouse.Sw  HKDDNPSLPP FERPEAEAMC TSFKENPTTF MGHYLHEVAR RHPYFYAPEL

                    101                                                150
Albu_Bovin.Sw  VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
Albu_Chick.Sw  ICQVEKLRDS YGAMADCCSK ADPERNECFL SFKVSQPDFV QPYQRPASDV
Albu_Horse.Sw  VATLRATYGE LADCCEKQEP ERNECFLTHK DDHPNLPKLK PEPDAQCAAF
Albu_Human.Sw  VATLRETYGE MADCCAKQEP ERNECFLQHK DDNPNLPRLV RPEVDVMCTA
Albu_Mouse.Sw  LYYAEQYNEI LTQCCAEADK ESCLTPKLDG VKEKALVSSV RQRMKCSSMQ

                    151                                                200
Albu_Bovin.Sw  KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQD CCQAEDKGAC
Albu_Chick.Sw  ICQEYQDNRV SFLGHFIYSV ARRHPFLYAP AILSFAVDFE HALQSCCKES
Albu_Horse.Sw  QEDPDKFLGK YLYEVARRHP YFYGPELLFH AEEYKADFTE CCPADDKLAC
Albu_Human.Sw  FHDNEETFLK KYLYEIARRH PYFYAPELLF FAKRYKAAFT ECCQAADKAA
Albu_Mouse.Sw  KFGERAFKAW AVARLSQTFP NADFAEITKL ATDLTKVNKE CCHGDLLECA

                    201                                                250
Albu_Bovin.Sw  LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
Albu_Chick.Sw  DVGACLDTKE IVMREKAKGV SVKQQYFCGI LKQFGDRVFQ ARQLIYLSQK
Albu_Horse.Sw  LIPKLDALKE RILLSSAKER LKCSSFQNFG ERAVKAWSVA RLSQKFPKAD
Albu_Human.Sw  CLLPKLDELR DEGKASSAKQ RLKCASLQKF GERAFKAWAV ARLSQRFPKA
Albu_Mouse.Sw  DDRAELAKYM CENQATISSK LQTCCDKPLL KKAHCLSEVE HDTMPADLPA

                    251                                                300
Albu_Bovin.Sw  FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
Albu_Chick.Sw  YPKAPFSEVS KFVHDSIGVH KECCEGDMVE CMDDMARMMS NLCSQQDVFS
Albu_Horse.Sw  FAEVSKIVTD LTKVHKECCH GDLLECADDR ADLAKYICEH QDSISGKLKA
Albu_Human.Sw  EFAEVSKLVT DLTKVHTECC HGDLLECADD RADLAKYICE NQDSISSKLK
Albu_Mouse.Sw  IAADFVEDQE VCKNYAEAKD VFLGTFLYEY SRRHPDYSVS LLLRLAKKYE

                    301                                                350
Albu_Bovin.Sw  CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
Albu_Chick.Sw  GKIKDCCEKP IVERSQCIME AEFDEKPADL PSLVEKYIED KEVCKSFEAG
Albu_Horse.Sw  CCDKPLLQKS HCIAEVKEDD LPSDLPALAA DFAEDKEICK HYKDAKDVFL
Albu_Human.Sw  ECCEKPLLEK SHCIAEVEND EMPADLPSLA ADFVESKDVC KNYAEAKDVF
Albu_Mouse.Sw  ATLEKCCAEA NPPACYGTVL AEFQPLVEEP KNLVKTNCDL YEKLGEYGFQ

                    351                                                400
Albu_Bovin.Sw  GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
Albu_Chick.Sw  HDAFMAEFVY EYSRRHPEFS IQLIMRIAKG YESLLEKCCK TDNPAECYAN
Albu_Horse.Sw  GTFLYEYSRR HPDYSVSLLL RIAKTYEATL EKCCAEADPP ACYRTVFDQF
Albu_Human.Sw  LGMFLYEYAR RHPDYSVVLL LRLAKTYETT LEKCCAAADP HECYAKVFDE
Albu_Mouse.Sw  NAILVRYTQK APQVSTPTLV EAARNLGRVG TKCCTLPEDQ RLPCVEDYLS
```

Figure 2.1.5 Multiple alignment of mammalian blood proteins using PILEUP. The first 400 residues of five high-scoring sequences identified from FASTA analysis (Fig. 2.1.4) are aligned.

Similarly, the program PIMA (Smith and Smith, 1992) uses an algorithm that takes advantage of secondary structure features to determine gap penalties during the multiple alignment process. Depiereux and Feytmans (1991) have used simultaneous multivariate multiple alignment methods to reveal a correspondence between physiochemical profiles and structurally conserved regions.

The multiple alignment programs described above take advantage of pairwise comparison algorithms and are subject to their limitations. Alignment algorithms that use a rigorous consideration of more than three sequences at once are computationally infeasible for most labs. Although some speed improvements have been obtained for the optimal alignment of three sequences simultaneously (Murata et al., 1985; Gotoh, 1986), new methods have to be developed for multiple sequence alignment. The earliest methods were sequence editors for interactive alignments. See Pearson and Miller

BCM MEDICAL CENTER LIBRARY

(1992) for a discussion of considerably improved dynamic programming algorithms for sequence analysis.

Multiple sequence alignment tends to be order-sensitive. A sequence of pairwise comparisons may not lead to the best overall alignment for a given set of sequences because the sequences are not independent (States and Boguski, 1990)—that is, they are not necessarily equidistant from a common ancestor.

The **consensus sequence** approach to multiple sequence alignment (Waterman, 1989, 1990; Waterman and Jones, 1990; Day and McMorris, 1993) is based on the construction of a single sequence that can be used to align all other given sequences; the sequence is determined by an iterative process of alignment followed by consensus sequence modification. A consensus residue can be stringently defined from a multiple alignment when the content of a column is represented by a single amino acid at that position.

There are many multiple alignment algorithms—too many to describe here. A list of available sequence similarity and alignment software can be accessed via the Internet as described in the section below on Internet Resources.

Once sequences have been aligned, the quality of the alignment should be assessed. It is clearly important to distinguish between chance similarities and true sequence similarities. One method for estimating the quality or **statistical significance** of an alignment (Doolittle, 1986) is to compare the alignment score with the mean score of a set of randomized sequences that are permutations of the original sequences and which preserve their length and relative composition. The further the alignment score is from the mean of the randomized set, the more confident we can be in the alignment. A possible objection to this data shuffling or permutation method is that the permuted sequences are not like real protein sequences (i.e., permuted sequences are unlikely to fold like real proteins).

Other methods (States and Boguski, 1990) include the following: McLachlan's (1972) "double matching probability" function, which is used to compute the frequency scores in infinite random sequences of the same composition as the given sequences; Reich and Meiske's stochastic random sequence model for determining the significance of window scores in dot-matrix analysis (Reich and Meiske, 1987); and Argos' significance tests (Argos and Vingron, 1990), based on actual protein sequences instead of the randomized sequences used by McLachlan (1972). Despite efforts to determine the quality of alignments, it remains true that low statistical significance does not always imply a lack of biological significance (Collins and Coulson, 1990).

Another important aspect of alignments is the **accuracy** of the original molecular sequence (States, 1992). Errors can be introduced in molecular sequence data in random or biased ways. The translation of nucleic acid sequences into amino acid sequences is very sensitive to indel errors: an error rate of 1% can lead to a frameshifting error in reading frames longer than 24 amino acids. Empirical estimates of sequence accuracy in current databases have revealed a rate of 3 to 5 errors per 1000 bases.

CLUSTER METHODS AND TREES

Once a multiple alignment has been performed and regions of common similarity have been identified among the proteins, the question of evolutionary distance will arise. There are several programs for determining the level of divergence between a group of sequences. Some are available over the Internet via the World Wide Web (WWW; see also Internet Resources section). The programs DISTANCES and GROWTREE from the GCG package (Genetics Computer Group, 1994) calculate pairwise distance measures and plot a dendrogram of those measures. The programs depend on an accurate multiple alignment of the sequences involved. This means that the multiple alignments should be examined carefully and then edited as necessary to incorporate additional knowledge that might be available about the structure or function of the proteins. Figure 2.1.6 shows the table of pairwise phylogenetic distances calculated by DISTANCES for the top nine albumins identified from the FASTA analysis (Fig. 2.1.4), and the phylogram produced by GROWTREE is shown in Figure 2.1.7.

The purpose of protein sequence comparison is to learn more about the structure, function, and evolution of proteins. It is known that different proteins and different parts of proteins change at characteristic but unequal rates in the course of evolution (Doolittle, 1986). Binding sites and catalytic units are conserved independently of their position, whereas surface regions change faster than buried regions. In addition, entire domains of a genome are subject to a number of editing operations such as domain joining, swapping, and reordering (States and Boguski, 1990). Thus, homologies

BU IASTATE UNIVERSITY LIBRARY

Matrix 1, dimension: 9

Key for column and row indices:

1 Albu_Bovin.Sw
2 Albu_Chick.Sw
3 Albu_Horse.Sw
4 Albu_Human.Sw
5 Albu_Mouse.Sw
6 Albu_Pig.Sw
7 Albu_Ranca.Sw
8 Albu_Rat.Sw
9 Albu_Sheep.Sw

	1	2	3	4	5	6	7	8	9
1	0.00	999.99	31.75	192.01	999.99	999.99	999.99	237.70	8.19
2		0.00	999.99	999.99	999.99	999.99	999.99	999.99	999.99
3			0.00	190.56	999.99	999.99	999.99	222.61	29.05
4				0.00	999.99	999.99	999.99	32.75	193.48
5					0.00	999.99	999.99	999.99	999.99
6						0.00	999.99	999.99	999.99
7							0.00	999.99	999.99
8								0.00	237.70
9									0.00

Figure 2.1.6 Evolutionary distance table for mammalian blood proteins calculated using the GCG DISTANCES program. Corrections to the calculations were done using the Kimura protein method as described in the GCG manual (Genetics Computer Group, 1994). The matrix values represent the number of matches between each pair of sequences divided by its length (999.9 = identity). These data were used to generate the GROWTREE phylogram presented in Figure 2.1.7.

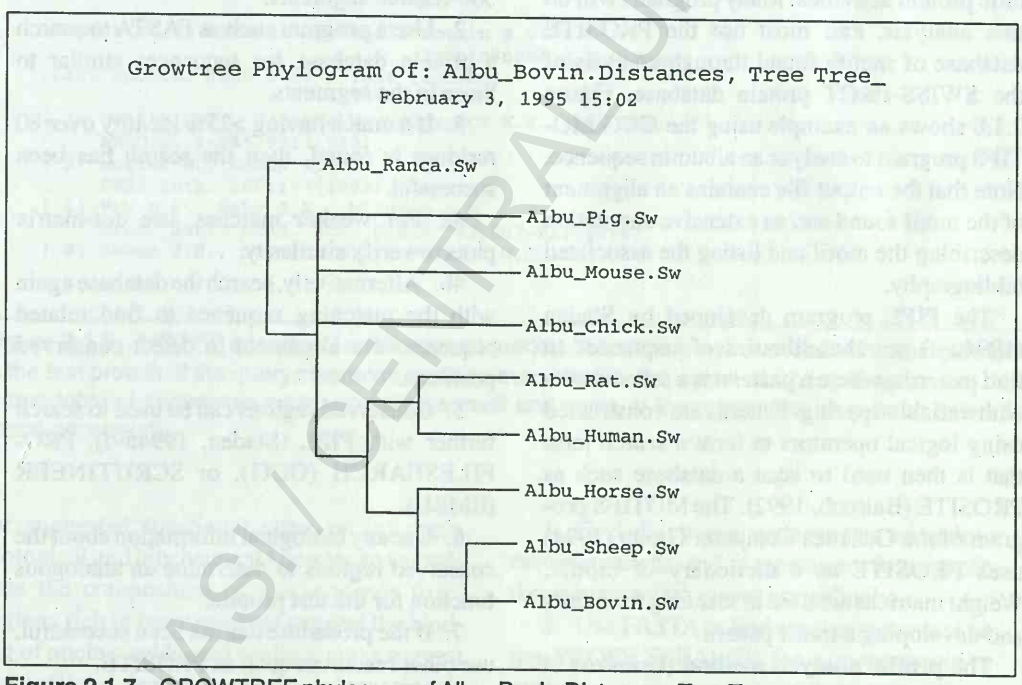


Figure 2.1.7 GROWTREE phylogram of Albu_Bovin.Distances, Tree Tree_1. Results of graphical cluster analysis of the distance data presented in Figure 2.1.6 are shown. In this representation sequences with the least divergence have the shortest branches.

BCU LASI / CENTRAL UNIVERSITY LIBRARY

might be found between entire proteins or just between certain parts of proteins.

IDENTIFICATION OF FUNCTIONAL SITES

Sequence comparisons are used to test for homology and to find functional sites and motifs. The goal is to establish structure and function for parts of an unknown sequence using known homologous or analogous (i.e., similar, but not homologous) sequences as sources of additional information. The goal is to obtain answers to questions of what functions the protein has and what its structure is. There are several paths the analysis can take.

One approach that will reveal much about potential functional sites in the protein is to look for motifs. **Motifs** are defined in three ways: as consensus patterns of amino acids that correspond to regions of proteins containing the common residues; as the excised alignment of several sequences that contain the motif for use as a database discriminator; or as a single representative sequence (States and Boguski, 1990; Parry-Smith and Atwood, 1992). Motif searches are usually done using a program that takes advantage of a dictionary of sequence patterns which have been identified with specific protein activities. Many programs will do this analysis, and most use the PROSITE database of motifs found through analysis of the SWISS-PROT protein database. Figure 2.1.8 shows an example using the GCG MOTIFS program to analyze an albumin sequence. Note that the output file contains an alignment of the motif found and an extensive annotation describing the motif and listing the associated bibliography.

The PIPL program developed by Staden (1994a-d) searches libraries of sequences to find patterns, where a **pattern** is a set of motifs with variable spacing. Patterns are constructed using logical operators to form a search term that is then used to scan a database such as PROSITE (Bairoch, 1992). The MOTIFS program of the Genetics Computer Group (1994) uses PROSITE as a dictionary of motifs. Weight matrices are used in scanning for motifs and developing a motif pattern.

The **profile analysis method** (Gribskov et al., 1988, 1990; Gribskov, 1994) also involves weight matrices and is related to the consensus method. It is used for sensitive similarity matching, for homology seeking, and for scanning databases for motifs. It measures the similarity between a given sequence, called the **target**, and a group of aligned sequences called the **probe**. The alignment of the probe uses a modified MDM_{78} mutational matrix. The matrix values of the profile are computed column-wise as the weighted average of the score for each residue at that position in the aligned sequences and the residue represented by the column of the profile; a row of the profile is the weighted average of the rows in the MDM_{78} table corresponding to the aligned residues at that position. The result is the set of properties that tend to be conserved at particular positions along the sequence. For detailed information on using profile analysis software to scan for homologies and motifs, see Gribskov (1994). A relatively new homologous sequence retrieval program (using the BLOSUM250 matrix) is dFLASH, which is available by E-mail server (Rigoutsos and Califano, 1994).

There are many ways to proceed in computational analysis of new sequences. Each will provide a wealth of information to guide further laboratory experimentation. The following procedures illustrate two approaches to protein sequence analysis and can be adapted as required.

Hodgman (1992) suggests the following protocol for protein function determination:

1. Split sequences into overlapping 200- to 300-residue segments.
2. Use a program such as FASTA to search a protein database for sequences similar to those in the segments.
3. If a match having >25% identity over 80 residues is found, then the search has been successful.

4a. For weaker matches, use dot-matrix plots to verify similarity.

4b. Alternatively, search the database again with the matching sequence to find related sequences for alignment to detect conserved positions.

5. Conserved regions can be used to search further with PIPL (Staden, 1994a-d), PROFILESEARCH (GCG), or SCRUTINEER (EMBL).
6. Use any biological information about the conserved regions to determine an analogous function for the test protein.
7. If the procedure has not been successful, use other resources such as PROSITE.
8. Check that all results are biologically credible.
9. If all else fails, (a) look for repeats within the sequence itself: if there are only two or three, it usually indicates that the protein has a repeated functional domain; (b) examine the distribution of particular amino acids or classes

```
Motifs Analysis

Atp_Gtp_A                (A,G)x4GK(S,T)     the motif from the Prosite database
                          (G)x{4}GK(S)      the pattern matched in this sequence
                 425: HIVLS GESYSGKS TNARL   the acutal region of sequence
                                             containing the motif

Attached annotation
****************************************
* ATP/GTP-binding site motif A (P-loop) *
****************************************

From sequence comparisons and crystallographic data analysis it has been shown
[1,2,3,4,5] that an appreciable proportion of proteins that bind ATP or GTP share
a number of more or less conserved sequence motifs.  The best conserved of these
motifs is a glycine-rich region, which probably forms a flexible loop between a beta-
strand and an alpha-helix.  This loop interacts with one of the phosphate groups of
the nucleotide.  This sequence motif is generally referred to as the 'A' consensus
sequence [1] or the 'P-loop' [5].

There are numerous ATP- or GTP-binding proteins in which the P-loop is found.We list
below a number of protein families for which the relevance of the presence of
such motif has been noted:

A list of proteins containing the motif has been omitted for brevity.

Not all ATP- or GTP-binding proteins are picked-up by this motif.  A number of
proteins escape detection because the structure of their ATP-binding site is
completely different from that of the P-loop.  Examples of such proteins are the
E1-E2 ATPases or the glycolytic kinases.  In other ATP- or GTP-binding proteins
the flexible loop exists in a slightly different form; this is the case for
tubulins or protein kinases.  A special mention must be reserved for adenylate
kinase, in which there is a single deviation from the P-loop pattern: in
the last position Gly is found instead of Ser or Thr.

-Consensus pattern: [AG]-x(4)-G-K-[ST]
-Sequences known to belong to this class detected by the pattern: a majority.
-Other sequence(s) detected in SWISS-PROT: in addition to the proteins listed above,
the 'A' motif is also found in a number of other proteins. Most of these proteins
probably bind a nucleotide, but others are definitively not ATP- or GTP-binding
(as for example chymotrypsin, or human ferritin light chain).

-Expert(s) to contact by email: Koonin E.V.
                                koonin@ncbi.nlm.nih.gov
-Last update: June 1994 / Text revised.

[ 1] Walker J.E., Saraste M., Runswick M.J., Gay N.J.
     EMBO J. 1:945-951(1982).
[ 2] Moller W., Amons R.
     FEBS Lett. 186:1-7(1985).
[ 3] Fry D.C., Kuby S.A., Mildvan A.S.
     Proc. Natl. Acad. Sci. U.S.A. 83:907-911(1986).
[ 4] Dever T.E., Glynias M.J., Merrick W.C.
```

Figure 2.1.8 MOTIFS analysis of an albumin sequence. This example shows results for a single motif found in the test protein. If the query sequence contains many motifs, the output can be quite extensive. Note that a long table of sequences which contain the motif and some of the references have been deleted due to space constraints.

for suggested functional sites; or (c) apply biological and biochemical knowledge to evaluate the composition of the sequence (e.g., regions rich in basic residues suggest the binding of nucleic acids, and acidic regions suggest generalized protein-protein interactions). Structure determination can begin with the Chou-Fasman (Chou and Fasman, 1974) or Garnier-Osguthorpe-Robson algorithm (Garnier et al., 1978); both are discussed in *UNIT 2.3*.

States and Boguski (1990) offer another scheme for the analysis of protein sequences:

1. Find all intrasequence repeats using dot-matrix methods, and, if repeats are found, split the sequence into pieces accordingly.
2. Use FASTA to find similar sequences or use PROFILESEARCH for a more sensitive search.
3. Compare the sequence to known motifs, for example, in the PROSITE database.
4. Search for hydrophobic segments using the Kyte-Doolittle (1982) method and for α helices and β strands using the hydrophobic moment (see *UNITS 2.2 & 2.3*).

BOTANY / CENTRAL UNIVERSITY LIBRARY

5. If step 1 or 2 was successful, then search for helices using the variation moment.

6. If step 1 or 2 was successful, calculate the hydrophobic correlation to estimate similarity.

7. Keeping in mind that the results will be only 55% to 65% accurate, compare the outputs of several secondary structure prediction methods with information from all available homologous sequences.

8. For antibodies, identify possible antigenic regions using hydrophilicity plots and antigenicity index plots.

9. Look for unusual features such as leucine zippers; this can be done with any program, such as PROSITE, that takes advantage of a dictionary of motifs.

10. Determine potential sites of cleavage by various chemicals; cleavage sites can be predicted by computer methods and verified experimentally.

INTERNET RESOURCES

The Internet has become an important resource in bioinformatics. There are new tools that make it easy to locate the latest software and databases. Since 1994 a new Internet-based network of hypertext tools that takes advantage of the World Wide Web (WWW) has emerged. To enter the WWW, you will need a network hypertext browser such as MOSAIC. Versions of MOSAIC, which was originally developed by the National Center for Supercomputer Applications (NCSA), are now freely available for most commonly used laboratory computers. If your local computer is connected to the Internet and your system has file transfer protocol (FTP) capability, connect to ftp.ncsa.uiuc.edu and log in under the username anonymous to retrieve a copy of MOSAIC for the computer you use most often. Once you have MOSAIC or another WWW browser running, we recommend connecting to the Australian National University's Bioinformatics server (http://life.anu.edu.au) and selecting "Molecular biology" from the menu. Table 2.1.1 reproduces the list of resources available. Another current resource of hyperlinks is available through the author using MOSAIC or Netscape at the George Mason Computational Biology Resource (http://www.science.gmu.edu/~michaels/comp_bio).

As indicated by that list, there is a tremendous number of resources available on the Internet. In addition, there are many useful text and program files available by FTP that are of interest to molecular biologists and other scientific and medical researchers. A few are presented here. For a listing of molecular sequence and related software, get the *Catalogue of Molecular Biology* from Genethon by connecting via ftp ftp.genethon.fr and entering cd pub/resig/catalogu then mget bio.catal*.

Amos Bairoch's bibliographic database on sequence analysis (*seqanalref*), the Listing of Molecular Biology Databases (*LiMB*), a directory of researchers who apply artificial intelligence (AI) techniques to problems in molecular biology (*aimbdb*), and many other valuable files can be found using anonymous FTP to ncbi.nlm.nih.gov. Bairoch's lists of molecular biology e-mail, FTP, and BBO servers are also available using anonymous FTP: type ftp expasy.hcuge.ch, then cd databases/info, then mget serv_ema.txt serv_ftp.txt serv_bbo.txt.

For a copy of the *BIOSCI FAQ*, a collection of Frequently Asked Questions, use ftp net.bio.net, then cd pub/BIOSCI/doc and get biosci.FAQ.

LITERATURE CITED

Allison, L., Wallace, C.S., and Yee, C.N. 1992. Finite-state models in the alignment of macromolecules. *J. Mol. Evol.* 35:77-89.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

Argos, P. 1987. A sensitive procedure to compare amino acid sequences. *J. Mol. Biol.* 193:385-396.

Argos, P. and Vingron, M. 1990. Sensitivity comparison of protein amino acid sequences. *Methods Enzymol.* 183:352-365.

Bairoch, A. 1992. PROSITE: A dictionary of protein sites and patterns. *Nucl. Acids Res.* 19:2241-2245.

Barton, G.J. 1990. Protein multiple sequence alignment and flexible pattern matching. *Methods Enzymol.* 183:403-428.

Barton, G.J. 1993. An efficient algorithm to locate all locally optimal alignments between two sequences allowing for gaps. *CABIOS* 9:729-734.

Brendel, V. 1992. PROSET—A fast procedure to create nonredundant sets of protein sequences. *Math. Comput. Modeling* 16(6/7):37-43.

Brendel, V., Bucher, P., Nourbaksh, I.R., Blaisdell, B.E., and Karlin, S. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. U.S.A.* 89:2002-2006.

Burks, C. 1990. The flow of nucleotide sequence data into data banks: Role and impact of large-scale sequencing projects. *In* Computers and DNA, Santa Fe Institute (G. Bell and T. Marr, eds.) pp. 35-45. Addison-Wesley, Reading, Mass.

Table 2.1.1 Australian National University (ANU) Molecular Biology Resources

Databases

French WWW genomic service, including *Caenorhabditis elegans* Database
BLAST Database Searches at NCBI
CompoundKB database—981 metabolic intermediate compounds
Codon usage tables (major species) & EMBL mirror
DNA Data Bank of Japan
EC enzyme database
EMBL data (Heidelberg)
ESTDB—Expressed Sequence Tag Database (TIGR)
FASTA Database search program (Virginia)
Genbank searches (Indiana)
Genbank/Swiss-Prot/Protein/PIR (via NIH)
GENETHON Human Genome Centre
Mendelian Inheritance in Man (index)
Metabolic intermediate compounds
Microbial germplasm
Miscellaneous, e.g. codon usage, profiles (Weizmann)
PIR (Houston)
Prosite (via NIH)
Protein databank (Brookhaven)
REF52 2D Gel Protein Database
REBASE—Restrict. Enzyme Data Base (NEB)
REBASE restriction enzymes
Swiss-Prot: EMBNet (Heidelberg) and ExPASy (Geneva)
SWISS-2DPAGE—Two-dimensional Polyacrylamide Gel Electrophoresis Database (Geneva)
Various (ICGEB, Italy)
Biologists's Control Panel

Bibliographies/Tutorials

Biosequence Comparison
Sequence analysis (search index)
Biocomputing bibliography
Biological Journals—current titles
Molecular biology journals
References to molecular biology algorithms
Periodical references to journals in molecular biology

Software

ANU software (gopher link)
SIMPLE34—Detection of sequence repetition in nucleic acid
MELANIE software packages for 2D PAGE computer analysis (Geneva)
NCBI repository
Sequence Analysis Tools (Trieste)
IUBio Biology Software and Database Archive (Indiana)
Quest II software for 2D gel proteins
CODA—Conservation Options and Decision Analysis
RAPDistance (gopher link)
RAPDistance (WWW link)

News groups

BioNet—various biological topics
The Scientist (biweekly news with a biotech slant)
BIOSCI mailing lists and newsgroup archives

Special topics

The Australian National Genomic Information Service (ANGIS)
Human Genome Mapping Project (UK)
Artificial Intelligence in molecular biology
Database of molecular biologists working in AI
Agricultural Genome World Wide Web server
A *Caenorhabditis elegans* DataBase (ACeDB)

Other sites

Australian National Genomic Information Service (ANGIS)
CAMIS Centre for Advanced Medical Informatics (Stanford)
Caenorhabditis Genetics Centre (CGC)
Chlamydomonas Genetics (Duke)
EMBNet
ExPASy (Geneva)
Harvard Biological Laboratories
Human Genome Center at Lawrence Berkeley Laboratory
Johns Hopkins University, BioInformatics
National Center for Biotechnology Information (NCBI)
QUEST Protein Database Center
Ribosomal Database Project (Argon Natl. Laboratory)
Molecular biology archives (list)
NIH GenoBase server
Othersites (gopher links)
Yeast Genome Information Server (Stanford)
WAIS—EC enzymes
WAIS—Molecular Biology
WAIS—indexes sorted alphabetically (ADFA)

BIOLOGY / CENTRAL UNIVERSITY LIBRARY

Chou, P.Y. and Fasman, G.D. 1974. Prediction of protein conformation. *Biochemistry* 13:222-244.

Collado-Vides, J. 1991. The search for a grammatical theory of gene regulation is formally justified by showing the inadequacy of context-free grammars. *CABIOS* 7:321-326.

Collins, J.F. and Coulson, A.F.W. 1990. Significance of protein sequence similarities. *Methods Enzymol.* 183:474-487.

Day, W.H.E. and McMorris, F.R. 1993. A consensus program for molecular sequences. *CABIOS* 9:653-656.

Dayhoff, M.O. 1978. Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Washington, D.C.

Depiereux, E. and Feytmans, E. 1991. Simultaneous and multivariate alignment of protein sequences: Correspondence between physicochemical profiles and structurally conserved regions (SCR). *Protein Eng.* 4:603-613.

De Rijk, P. and De Wachter, R. 1993. DCSE, an interactive tool for sequence alignment and secondary structure search. *CABIOS* 9:735-740.

Doolittle, R.F. 1981. Similar amino acid sequences: Chance or common ancestry? *Science* 214:167-339.

Doolittle, R.F. 1986. Of URFs and ORFs: A Primer on How to Analyze Derived Amino Acid Sequences. University Science Books, Ann Arbor, Mich.

Doolittle, R.F. 1989. Redundancies in protein sequences. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 599-623. Plenum, New York.

Doolittle, R.F. 1990. What we have learned and will learn from sequence databases. *In* Computers and DNA, Santa Fe Institute (G. Bell and T. Marr, eds.) pp. 21-31. Addison-Wesley, Reading, Mass.

Dumas, J.P. and Nunio, J. 1982. Efficient algorithm for folding and comparing nucleic acid sequences. *Nucl. Acids Res.* 10:197-206.

Eroshkin, A.M., Zhilkin, P.A., and Fomin, V.I. 1993. Algorithm and computer program: Pro_Anal for analysis of relationship between structure and activity in a family of proteins or peptides. *CABIOS* 9:491-497.

Fitch, W.M. 1966. An improved method of testing for evolutionary homology. *J. Mol. Biol.* 16:9-16.

Fitch, W.M. 1969. Locating gaps in amino acid sequences to optimize the homology between two proteins. *Biochem. Genet.* 3:99-108.

Fuchs, R. 1994. Fast protein block searches. *CABIOS* 10:79-80.

Garnier, J., Osguthorpe, D.J., and Robson, B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97-120.

Genetics Computer Group. 1994. GCG Program Manual for the Wisconsin Package, Version 8, September 1994. Genetics Computer Group Inc., Madison, Wis.

George, D., Hunt, L.T., and Barker, W.C. 1990. Mutation data matrix and its uses. *Methods Enzymol.* 183:333-351.

Gibbs, A.J. and McIntyre, G.A. 1970. The diagram, a method for comparing sequences. *J. Biochem.* 16:1-11.

Gotoh, O. 1986. Alignment of three biological sequences with an efficient traceback procedure. *J. Theor. Biol.* 121:327-337.

Gribskov, M., Homyak, M., Edenfield, J., and Eisenberg, D. 1988. Profile scanning for three-dimensional structural patterns in protein sequences. *CABIOS* 4:61-66.

Gribskov, M., Luethy, R., and Eisenberg, D. 1990. Profile analysis. *Methods Enzymol.* 183:146-159.

Gribskov, M. 1994. Profile analysis. Computer analysis of sequence data. *Methods Mol. Biol.* 24:247-266.

Henikoff, S. and Henikoff, J.G. 1993. *Proteins Struct. Funct. Genet.* 17:49-61.

Heringa, J., Sommerfeldt, H., Higgins, D.G., and Argos, P. 1992. OBSTRUCT: A program to obtain the largest cliques from a protein sequence set according to structural resolution and sequence similarity. *CABIOS* 8:599-600.

Hodgman, T.C. 1992. Nucleic acid and protein sequence management. *In* Microcomputers in Biochemistry: A Practical Approach (C.F.A. Bryce, ed.) pp. 131-158. IRL Press, Oxford.

Kanaoka, M., Kishimoto, F., Ueki, Y., and Umeyama, H. 1989. Alignment of protein sequences using the hydrophobic core scores. *Protein Eng.* 2:347-351.

Karlin, S.P., Morris, M., Ghandour, G., and Leung M.-Y. 1988. Algorithms for identifying local molecular sequence features. *CABIOS* 4:41-51.

Karlin, S.P., Ost, F., and Blaisdell, B.E. 1989. Patterns in DNA and amino acid sequences and their statistical significance. *In* Mathematical Methods for DNA Sequences (M.S. Waterman, ed.) pp. 133-157. CRC Press, Boca Raton, Fla.

Karlin, S., Bucher, P., and Brendel, V. 1991. Statistical methods and insights for protein and DNA sequences. *Annu. Rev. Biophys. Chem.* 20:175-203.

Kruskal, J.B. 1983. An overview of sequence comparison. *In* Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison (D. Sankoff and J.B. Kruskal, eds.) pp. 1-44. Addison-Wesley, Reading, Mass.

Kruskal, J.B. and Sankoff, D. 1983. An anthology of algorithms and concepts for sequence comparison. *In* Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison (D. Sankoff and J.B. Kruskal, eds.) pp. 265-310. Addison-Wesley, Reading, Mass.

BU IAS LIBRARY

Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.

Landau, G.M., Vishkin, U., and Nussinov, R. 1988. Locating alignments with *k* differences for nucleotide and amino acid sequences. *CABIOS* 4:19-24.

Landau, G.M., Vishkin, U., and Nussinov, R. 1990. Fast alignment of DNA and protein sequences. *Methods Enzymol.* 183:487-502.

Landes, C., Henaut, A., and Risler, J.-L. 1993. Dotplot comparisons by multivariate analysis (DOCMA): A tool for classifying protein sequences. *CABIOS* 9:91-196.

Lipman, D.J. and Pearson, W.R. 1985. Rapid and sensitive protein similarity searches. *Science* 227:1435-1441.

Livingstone, C.D. and Barton, G.F. 1993. Protein sequence alignments: A strategy for the hierarchical analysis of residue conservation. *CABIOS* 9:745-756.

Maizel, J.V. and Lenk, R.P. 1981. Enhanced graphic matrix analysis of nucleic acids and protein sequences. *Proc. Natl. Acad. Sci. U.S.A.* 78:7665-7669.

McLachlan, A.D. 1971. Test for comparing related amino acid sequences: Cytochrome *c* and cytochrome *c*-551. *J. Mol. Biol.* 61:409-424.

McLachlan, A.D. 1972. Repeating sequences and gene duplication in proteins. *J. Mol. Biol.* 72:417-437.

Michaels, G.S., Taylor, R., Hagstrom, R., Price, M., and Overbeek, R. 1993. Searching for genomic organizational motifs: Explorations of the *E. coli* chromosome. *Comp. Chem.* 17:209-217.

Mrazek, J. and Kypr, J. 1993. UNIREP: A microcomputer program to find unique and repetitive nucleotide sequences in genomes. *CABIOS* 9:355-360.

Murata, M., Richardson, J.S., and Sussman, J.L. 1985. Simultaneous comparison of three protein sequences. *Proc. Natl. Acad. Sci. U.S.A.* 82:3073-3077.

Nedde, D.N. and Ward, M.O. 1993. Visualizing relationships between nucleic acid sequences using correlation images. *CABIOS* 9:331-335.

Needleman, S.B. and Wünsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443-453.

Panjukov, V.V. 1993. Finding steady alignments: Similarity and distance. *CABIOS* 9:285-290.

Parry-Smith, D.J. and Atwood, T.K. 1992. ADSP—A new package for computational sequence analysis. *CABIOS* 8:451-459.

Pearson, W.R. 1990. Rapid and sensitive comparison with FASTP and FASTA. *Methods Enzymol.* 183:63-98.

Pearson, W.R. 1994. Using the FASTA program to search protein and DNA sequence databases. *Methods Mol. Biol.* 24:365-389.

Pearson, W.R. and Miller, W. 1992. Dynamic programming algorithms for biological sequence comparison. *Methods Enzymol.* 210:576-610.

Pevzner, P.A. 1992. Statistical distance between texts and filtration methods in sequence comparison. *CABIOS* 8:121-127.

Pizzi, E.M., Attimonelli, M., Liuni, S., Frontali, C., and Saccone, C. 1991. A simple method for global sequence comparison. *Nucl. Acids Res.* 20:131-136.

Reich, J.G. and Meiske, W. 1987. A simple statistical significance test of window scores in large dot matrices obtained from protein or nucleic acid sequences. *Comput. Appl. Biosci.* 3:25-30.

Rigoutsos, I. and Califano, A. 1994. Searching in parallel for similar strings. *IEEE Computatl. Sci. Eng.* 60-75.

Robson, B. and Greaney, P.J. 1992. Natural sequence code representations for compression and rapid searching of human-genome-style databases. *CABIOS* 8:283-289.

Rohde, K. and Bork, P. 1993. A fast, sensitive pattern-matching approach for protein sequences. *CABIOS* 9:183-189.

Searls, D. 1993. The computational linguistics of biological sequences. *In* Artificial Intelligence and Molecular Biology (L. Hunter, ed.) pp. 47-120. MIT Press, Cambridge, Mass.

Sellers, P.H. 1974. On the theory and computation of evolutionary distances. *SIAM J. Appl. Math.* 26:787-793.

Smith, R.F. and Smith, T.F. 1992. Pattern-induced multisequence alignment (PIMA) algorithm employing secondary structure–dependent gap penalties for use in comparative protein modeling. *Protein Eng.* 5:35-41.

Smith, T.F. and Waterman, M.S. 1981. Comparative biosequence metrics. *J. Mol. Evol.* 18:38-46.

Staden, R. 1994a. Statistical and structural analysis of protein sequences. *Methods Mol. Biol.* 24:125-130.

Staden, R. 1994b. Searching for motifs in protein sequences. *Methods Mol. Biol.* 24:131-139.

Staden, R. 1994c. Using patterns to analyze protein sequences. *Methods Mol. Biol.* 24:141-154.

Staden, R. 1994d. Comparing sequences. *Methods Mol. Biol.* 24:155-170.

States, D.J. 1992. Molecular sequence accuracy: Analyzing imperfect data. *Trends Genet.* 8:52-55.

States, D.J. and Boguski, M.S. 1990. Sequence Analysis Primer. Stockton Press, New York.

Streletc, V.B., Shindyalov, I.N., Kolchanov, N.A., and Lim, H.A. 1991. Fast, statistically based alignment of amino acid sequences on the base of diagonal fragments of dot matrices. *CABIOS* 8:529-534.

Taylor, W.R. 1988. Pattern matching methods in protein sequence comparison and structure prediction. *Protein Eng.* 2:77-86.

EGYETEMI KÖZPONTI KÖNYVTÁR LIBRARY

Tyler, E., Horton, M.R., and Krause, P.R. 1991. A review of algorithms for molecular sequence comparison. *Computers Biomed. Res.* 24:72-96.

Waterman, M.S. 1989. Sequence alignments. *In* Mathematical Methods for DNA Sequences (M.S. Waterman, ed.) pp. 53-90. CRC Press, Boca Raton, Fla.

Waterman, M.S. 1990. Consensus patterns in sequences. *In* Mathematical Methods for DNA Sequences (M.S. Waterman, ed.) pp. 93-115. CRC Press, Boca Raton, Fla.

Waterman, M.S. and Eggert, M. 1991. A new algorithm for best subsequence alignments with application to tRNA-rRNA comparisons. *J. Mol. Biol.* 197:723-728.

Waterman, M.S. and Jones, R. 1990. Consensus methods for DNA and protein sequence alignment. *Methods Enzymol.* 183:221-237.

Wilbur, W.J. and Lipman, D.J. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. U.S.A.* 80:726-730.

Contributed by George Michaels and Robert Garian
George Mason University
Fairfax, Virginia

BSU IASI / CENTRAL UNIVERSITY LIBRARY

Hydrophobicity Profiles for Protein Sequence Analysis

INTRODUCTION

Seminal work by Anfinsen et al. (1961) showed that the three-dimensional structure of a protein is determined by its amino acid sequence. Thus, significant effort has been aimed at unraveling the complex relationship between sequence and structure. Although there is no immediate prospect of a general solution to the protein folding problem, stepwise empirical methods have been developed that analyze protein sequences and successfully provide predictive results concerning protein structure.

The work of Kauzmann (1959) determined that hydrophobic interactions are a major force in protein folding. Accounting for the hydrophobicity of amino acids within a protein is thus the first step toward understanding protein folding (Tanford, 1980). Hydrophobicity studies have resulted in a wide range of **hydropathy scales**, indices that attempt to quantify the relative hydrophobicity of the amino acids. The actual hydrophobic content attributed to each amino acid is dependent on the methodology used to estimate it and may vary from scale to scale. Some scales are based solely on physicochemical measurements such as the free energy of side-chain transfer (Nozaki and Tanford, 1971; Wolfenden et al., 1981). Other scales are based on statistical analysis of three-dimensional structures (Chothia, 1976; Rose et al., 1985). In addition, scales can be a combination of both procedures (Kyte and Doolittle, 1982). The general observation that hydrophobic residues tend to be buried in the interior of the protein whereas hydrophilic residues prefer to be exposed to solvent is predominant in all scales.

Generating a hydropathy profile requires two things: (1) the primary structure and (2) the hydropathy scale. The primary structure can be obtained by translating the cDNA or by sequencing the protein. Hydropathy profiles for a protein are obtained by averaging or combining the hydrophobicity or hydrophilicity of an individual amino acid (taken from a particular hydropathy scale) with the values of several neighboring residues along the protein sequence, then plotting the local averages or sums against the amino acid sequence. In such hydrophobicity plots, regions of hydrophobicity and hydrophilicity appear as maxima and minima, respectively. A fundamental use of hydropathy profiles, therefore, is predicting the overall hydropathy (hydrophobicity or hydrophilicity) of a protein from its amino acid sequence. The distribution of local hydropathy values can provide insight into the overall folding pattern of the protein (see Applications).

Numerous procedures have been developed for calculating the hydropathy profile of a protein based on the free energy of side-chain transfer from water or vacuum to an organic solvent (Nozaki and Tanford, 1971; Tanford, 1980; Wolfenden et al., 1981; Fauchere and Pliska, 1983). The procedures differ primarily in the hydropathy scale used, and the different hydropathy scales have arisen, in part, to address specific applications. For example, the hydrophilicity scale of Hopp and Woods (1981) was developed to locate antigenic segments; the Goldman, Engleman, and Steitz (GES) scale (Engleman et al., 1986) was developed to identify transmembrane regions in a protein.

Although individual scales were developed for different purposes or by alternative procedures, significant similarity exists in the resultant profiles. Most hydropathy plots show similar distributions of minima and maxima in the sequence profiles. In Figure 2.2.1, the result of applying the most widely used hydrophobicity scale, that of Kyte and Doolittle (1982), is compared to results of applying the scales of Hopp and Woods (1981) and Eisenberg (1984) to hen egg white lysozyme. The similarity in profiles is partially due to the fact that maxima and minima often reflect secondary structure elements; regions of hydrophobicity and hydrophilicity can be correlated with β sheets and β turns, respectively (Kuntz, 1972; Roy and Rose, 1980).

Hydropathy profiles can be used to examine the surface features of proteins in order to generate hypotheses that can be confirmed experimentally. For example, such analyses received widespread use in the prediction of antigenic and other interaction sites on proteins. Furthermore, when used in conjunction with other secondary structure prediction algorithms (see *UNIT 2.3*), hydropathy profiles have been shown to correlate well with α-helical, turn, and β-sheet propensities. As such, they have proved to be invaluable for predicting protein secondary structure. The most extensive use of hydropathy profiles has been for

Copyright © 1995 by John Wiley & Sons, Inc.

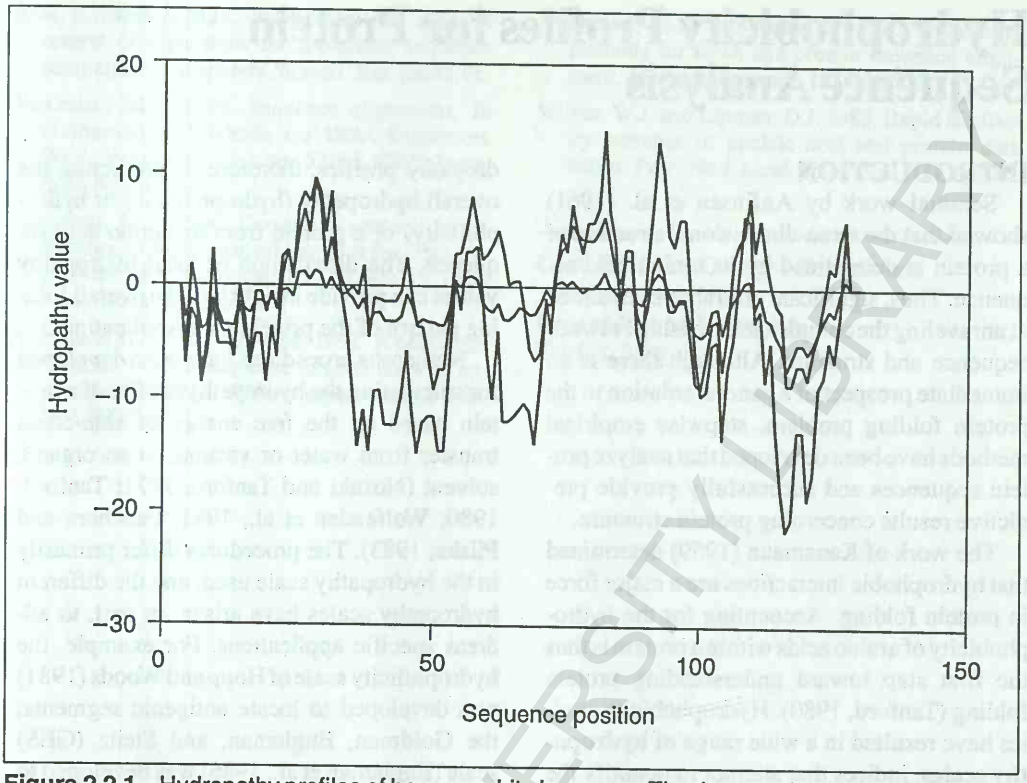


Figure 2.2.1 Hydropathy profiles of hen egg white lysozyme. The three profiles were created with the Kyte and Doolittle, the Hopp and Woods, and the Eisenberg scales using a scanning window of 7 amino acids. The Hopp and Woods scale was modified as described in the text to match other hydropathy scales.

deducing transmembrane regions in membrane proteins.

Previous reviews describe systematic evaluations of hydropathy procedures and provide criteria for selecting the most appropriate amino acid hydropathy scale (Engleman et al., 1986; Cornette et al., 1987; Banghan, 1988; Hopp, 1989; Rose and Dworkin, 1989). This unit describes the application of hydrophobicity plots to typical problems and provides suggested uses for a few selected scales.

METHODOLOGY

Selecting Hydropathy Scales

The first step in generating a hydropathy profile (or profiles) for an individual amino acid sequence is selecting the appropriate scale(s). Table 2.2.1 lists several widely used scales.

This unit focuses primarily on four hydropathy scales: the Hopp and Woods, the Kyte and Doolittle, the Eisenberg, and the GES scales. Differences among the hydropathy indices for each amino acid in the four scales (Table 2.2.2) arise primarily from differences in the experimental data from which the scales were derived and from the methodologies by which each scale was optimized. The hydrophilicity scale derived by Hopp and Woods is based on solubility values of hydrophobic and hydrophilic amino acids (Nozaki and Tanford, 1971; Levitt, 1976). In the Kyte and Doolittle and the Eisenberg scales, the hydropathy index is a measure of the relative affinity of an amino acid for hydrophobic phases.

The GES scale is based on physicochemical considerations of helical structure and several energetic factors that reflect partitioning of an amino acid side chain between aqueous solution and a membrane bilayer. Both the Kyte and Doolittle and the Eisenberg scales are combinations of earlier scales. The Kyte and Doolittle scale combines values from the water-to-vapor transfers and the internal/external distributions of amino acid residues from known structures as determined by Chothia (1976). The Eisenberg scale is a consensus of the Wolfenden (Wolfenden et al., 1981), Chothia (Chothia, 1976), Janin (Janin, 1979), and von Heijne (von Heijne and Blomberg, 1979) scales and is designed to minimize effects of the outlying values observed for some amino acids in previous scales.

BU IAS CENTRAL UNIVERSITY LIBRARY

Table 2.2.1 Hydropathy Scales

Reference scale	Literature	Derivation[a]
Bigelow	Bigelow, 1967	Experimental
Tanford	Nozaki and Tanford, 1971	Experimental
Chothia	Chothia, 1976	Theoretical
Levitt	Levitt, 1976	Experimental
Janin	Janin, 1979	Theoretical
von Heijne and Blomberg	von Heijne and Blomberg, 1979	Theoretical
Roy and Rose	Roy and Rose, 1980	Experimental
Hopp and Woods	Hopp and Woods, 1981	Experimental
Wolfenden	Wolfenden et al., 1981	Experimental
Argos	Argos et al., 1982	Theoretical
Fraga	Fraga, 1982	Theoretical
Kyte and Doolittle	Kyte and Doolittle, 1982	Combination
Fauchere and Pliska	Fauchere and Pliska, 1983	Experimental
Eisenberg	Eisenberg, 1984	Combination
Hopp	Hopp, 1984	Theoretical
Rose	Rose et al., 1985	Theoretical
GES	Engleman et al., 1986	Theoretical
Parker, Guo, and Hodges	Parker et al., 1986	Experimental
Esposti et al.	Esposti et al., 1990	Theoretical

[a]Experimental, largely original data from physicochemical experiments; theoretical, derivation based on analysis of structural data; combination, consensus of multiple scales.

There are several other interesting scales that deserve mention. The first hydropathy scale for membrane proteins was developed and revised by von Heijne (von Heijne and Blomberg, 1979) and corresponds rather closely to the GES scale described above (Engleman et al., 1986). A hydrophilicity scale by Parker, Guo, and Hodges (1986) provides new hydrophilicity values that are determined from HPLC retention times of model synthetic peptides. Another important set of scales includes the statistical scales developed by Argos and co-workers for prediction of membrane-buried helices (Argos et al., 1982; Rao and Argos, 1986). These scales are derived from the statistical preference of specific amino acid residues to form the integral regions in transmembrane proteins, and they have been shown to correlate with highly resolved crystal structures much better than the physicochemical hydropathy scales (Esposti et al., 1990).

Because all hydropathy scales are conceptually similar, the strategy of applying them to a given problem should be analogous regardless of the particular scale chosen. Moreover, although there may always be debates over which of the scales (methods) will prove to be best, experience has shown that application of several different hydropathy schemes and alternative predictive rules will tend to build a consensus picture of protein hydrophobicity.

For this unit the values of the Hopp and Woods and GES scales are multiplied by –1 so that these scales comply with others having peak maxima corresponding to hydrophobicity and peak minima corresponding to hydrophilicity (the scale values were not modified in Table 2.2.2). The Hopp and Woods and GES plots shown are therefore inverse plots.

Applying the Scanning Window Averaging Technique

After the appropriate scale(s) has been selected, the next stage is generating the hydropathy profile for a protein. To create the profile, the numerical values derived from the hydropathy scales for each residue in the protein are repetitively averaged over the length of protein sequence. This procedure is commonly referred to as the scanning window averaging technique (Hopp and Woods, 1981; Kyte and Doolittle, 1982). In practice, the hydropathy is obtained by summing the individual hydropathy values (hydrophobicities or hydrophilicities) for a given number of contiguous amino acids along the protein chain. The sum of

BOSTON UNIVERSITY LIBRARY

Table 2.2.2 Hydrophobicity Indices for Selected Scales

Amino acid	H & W[a]	K & D[b]	Eisenberg[c]	GES[d]
Neutral, hydrophobic, aliphatic				
Gly	0.0	−0.4	0.16	−1.0
Ala	−0.5	1.8	0.25	−1.6
Val	−1.5	4.2	0.54	−2.6
Ile	−1.8	4.5	0.73	−3.1
Leu	−1.8	3.8	0.53	−2.8
Met	−1.3	1.9	0.26	−3.4
Neutral, hydrophobic, aromatic				
Phe	−2.5	2.8	0.61	−3.7
Tyr	−2.3	−1.3	0.02	0.7
Trp	−3.4	−0.9	0.37	−1.9
Neutral, hydrophilic				
Ser	0.3	−0.8	−0.26	−0.6
Thr	−0.4	−0.7	−0.18	−1.2
Asn	0.2	−3.5	−0.64	4.8
Gln	0.2	−3.5	−0.69	4.1
Acidic, hydrophilic				
Asp	3.0	−3.5	−0.72	9.2
Glu	3.0	−3.5	−0.62	8.2
Basic, hydrophilic				
His	−0.5	−3.2	−0.40	3.0
Lys	3.0	−3.9	−1.1	8.8
Arg	3.0	−4.5	−1.8	12.3
Thiol-containing				
Cys	−1.0	2.5	0.04	−2.0
Imino acid				
Pro	0.0	−1.6	−0.07	0.2

[a]Hopp and Woods, 1981.
[b]Kyte and Doolittle, 1982.
[c]Eisenberg, 1984.
[d]Engleman et al., 1986.

hydropathy values is calculated progressively through the sequence, with each amino acid being the start of a new sum (Fig. 2.2.2). The resultant sums (or averages) are plotted versus the sequence position at the position of the central (or first) amino acid in the segment being summed. The choice of the center versus the first residue has no effect on the calculated profiles but shifts the sequence axis by half the window size. We prefer to plot versus the center position as it seems more intuitive. Similarly, the choice of summing or averaging has no effect on the profile but modifies the relative values plotted on the *y*-axis. We use sums for the profiles presented in this unit. This approach displays the general hydropathic character of a protein.

The ability to discern structurally interpretable features in the calculated hydropathy profile is critically dependent on the width of the scanning window, that is, the number of amino acids included in the summation. Figure 2.2.3 shows the effect of altering the size of the scanning window on hydropathy profiles generated for lysozyme. In Figure 2.2.3A, where the window size is set to 1 amino acid, the resultant hydropathy profile is featureless, providing essentially no structural information. On introducing a window size of 7 amino acids, well-resolved peaks appear in the hydropathy

BU IAS CENTRAL UNIVERSITY LIBRARY

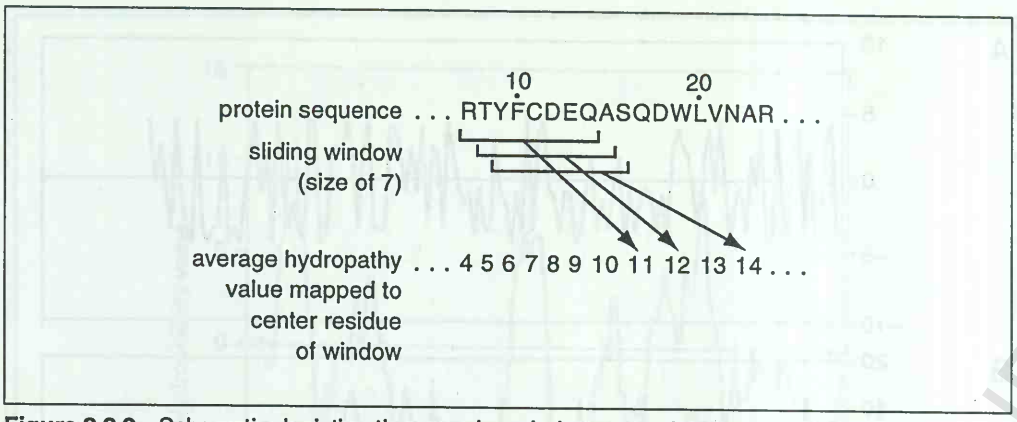

10
20
protein sequence . . . RTYFCDEQASQDWLVNAR . . .
sliding window
(size of 7)
average hydropathy
value mapped to
center residue
of window
. . . 4 5 6 7 8 9 10 11 12 13 14 . . .

Figure 2.2.2 Schematic depicting the scanning window averaging technique.

profile (Fig. 2.2.3B). Increasing the window size to 11 amino acids (Fig. 2.2.3C) results in a smoothing or broadening of observed peaks in the profile, and with window sizes of 15 or 19 amino acids (Fig. 2.2.3D,E), the profiles again have few well-resolved features.

From the family of profiles in Figure 2.2.3, it is readily apparent that window sizes exceeding 10 amino acids result in a loss of local information. This decrease in local information leads to increased errors in predicting hydrophobic regions in globular proteins (Kyte and Doolittle, 1982). In contrast, significant correlations do exist between hydropathy profiles generated with window sizes of 5 to 9 amino acids and known structural elements (Roy and Rose, 1980; Kyte and Doolittle, 1982; Rao and Argos, 1986; Cornette et al., 1987).

In selecting the appropriate window size for a particular problem, the length of the structural property being investigated should be used as a guide. For example, transmembrane helices are roughly 20 residues in length, but the helices of globular proteins tend to be much shorter. One might therefore opt for a larger window size if searching for transmembrane helices, whereas a length of 7 or 9 amino acids is more appropriate for predicting surface sites or secondary structure. In cases of uncertainty, analysis of profiles generated with various window sizes will afford the most reliable results.

Threshold or Cutoff Values

To facilitate the analysis of hydropathy profiles, it is convenient to define a **threshold value**, that is, the amplitude above which the hydropathy value must remain for a particular amino acid to be considered part of a contiguous region. Although somewhat arbitrary, the actual choice of threshold value for a particular profile is often derived empirically or from experience. For example, if one were interested in analyzing the sequences of known membrane proteins for transmembrane helices, then an appropriate threshold value might be one which delineates peaks in the hydropathy profile to approximately 20 amino acids (see section on Predicting Transmembrane Regions in Proteins). As with selecting scanning window size, analyzing profiles with different threshold values should yield more reliable results until experience is gained.

APPLICATIONS

Predicting Interaction Sites

Hydropathy profiles have found widespread use in investigations aimed at elucidating the antigenic structure and surface features of proteins for which there are no three-dimensional structures (Hopp and Woods, 1981; Krystek et al., 1985a,b). Antigenic sites tend to occur on protein surfaces, where predominantly hydrophilic residues are both exposed to solvent and accessible for interaction with antibody. One would therefore anticipate good correlation between antigenic sites and regions of high hydrophilicity. A test of this hypothesis carried out by Tanaka et al. (1985) reported a success rate of 56%.

Figure 2.2.4 depicts the Hopp and Woods profiles for myoglobin generated with a window size of 7 amino acids. Regions of hydrophilicity are detected as peak minima in the profile. The lines at the bottom of Figure 2.2.4 indicate the locations of five major antigenic sites determined experimentally (Atassi, 1984). All of the antigenic sites are associated with regions of high hydrophilicity. In addition, there are two regions of high hydrophilicity (peaks at positions 40 to 45 and 78 to 84) that are not associated with antigenicity. It is plau-

DUKE UNIVERSITY LIBRARY

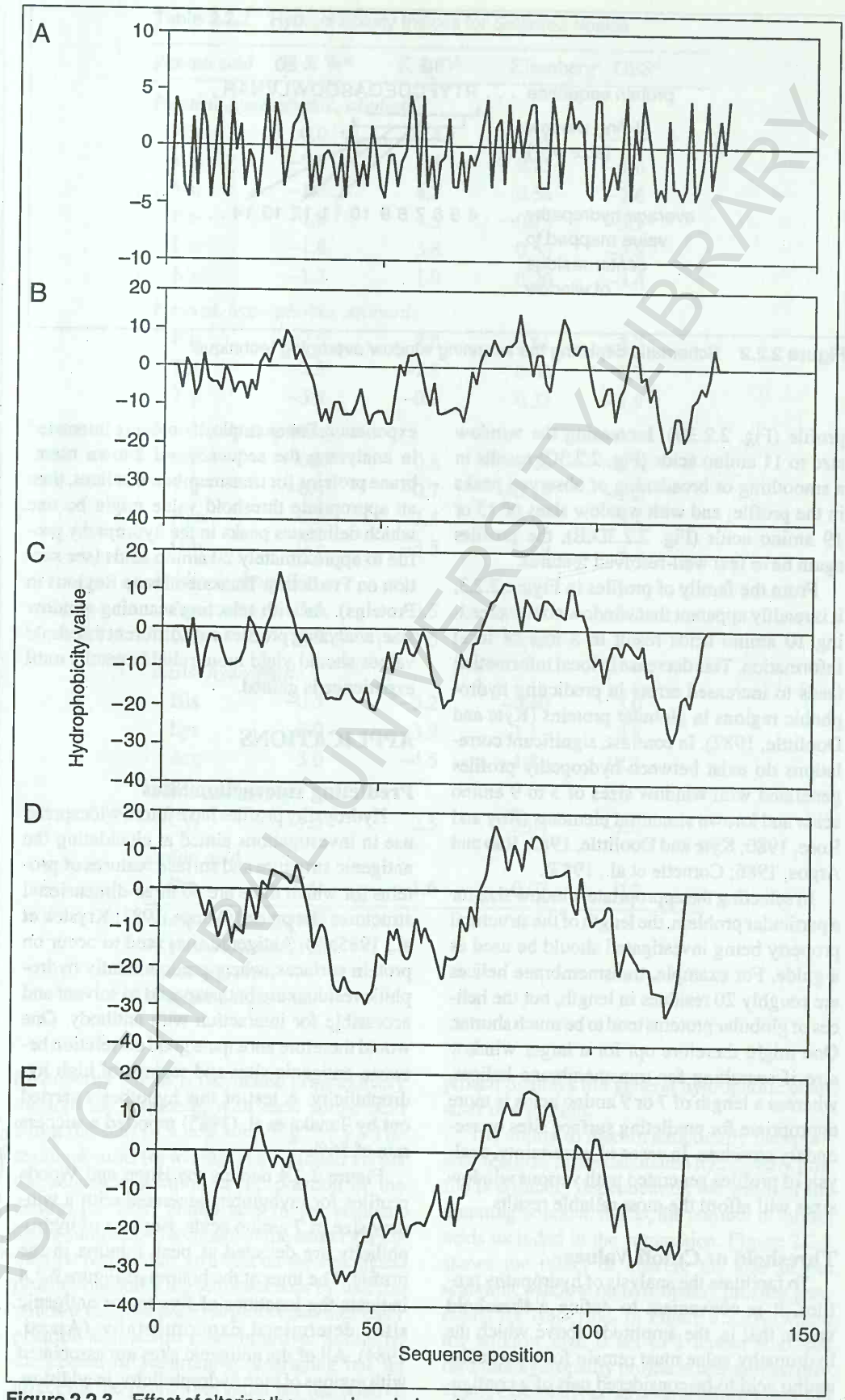


Figure 2.2.3 Effect of altering the scanning window size on the resolution of the hydropathy profile of lysozyme. All profiles were created with the Kyte and Doolittle scale. Window sizes were set to **(A)** 1, **(B)** 7, **(C)** 11, **(D)** 15, and **(E)** 19 amino acids. Note that the scale in **A** is reduced to show the profile more clearly.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

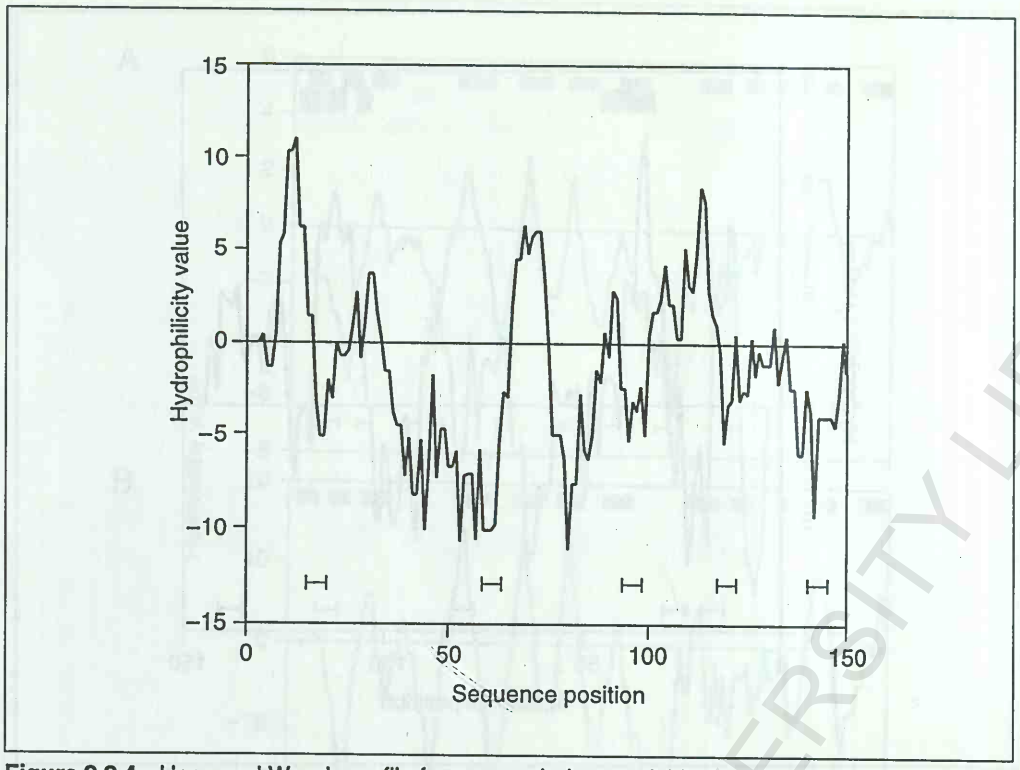


Figure 2.2.4 Hopp and Woods profile for sperm whale myoglobin using a scanning window of 7 amino acids. Lines below the profile indicate the locations of five experimentally determined antigenic sites (Atassi, 1984).

sible that these correspond to antigenic sites that have yet to be identified experimentally, or they may point to an alternate, functionally significant property of myoglobin. Whatever the case, these results serve to emphasize the point that hydropathy profiles should be used only as a guide for predicting antigenicity.

Additional features of protein surfaces can be highlighted by hydropathy profiles. As an example, consider human hemoglobin, whose three-dimensional structure is known (Fermi, 1975). Hemoglobin is found as a tetramer in its native state, being composed of two α- and two β-subunits. In the native protein the β- and α-subunits are folded against one another, forming contact regions that more closely resemble the interior of a globular protein. Because these contact regions are hydrophobic, one might examine hydropathy profiles for regions of hydrophobicity as a means of identifying potential subunit interacting surfaces.

The Hopp and Woods profile for human hemoglobin is shown in Figure 2.2.5. The regions of contact in the β-subunit of human hemoglobin (Yoshioka and Atassi, 1986) are indicated as bars above the profile. The two largest hydrophobic peaks correspond to regions of the β-subunit in contact with each of the α-subunits. Analogous analysis of hydropathy profiles has been successfully used to identify the subunit interacting surfaces for the α-subunits of glycoprotein hormones lutropin and follitropin (Krystek et al., 1991, 1992).

Human hemoglobin also provides a demonstration of the dangers of using hydropathy profiles for analyzing antigenic sites of proteins that possess significant quaternary structure. In Figure 2.2.5 the lines below the profile indicate the regions of five major continuous antigenic determinants. Although two sites correlate with regions of hydrophilicity, the others do not. In such cases the amino acid sequence alone is not sufficient for predictability. This may be due in part to the unique nature of proteins composed of multiple subunits.

Correlating Profiles with Secondary Structure

An increasingly important application of hydropathy profiles is their use in conjunction with secondary structure prediction tools (see *UNIT 2.3*). Figure 2.2.6 shows the hydropathy profiles for interleukin 1β (IL-1β) generated with the Eisenberg scale (Fig. 2.2.6A) and the Kyte and Doolittle scale (Fig. 2.2.6B), using a scan window of 7 amino acids. The solid bars

DALHOUSIE UNIVERSITY LIBRARY

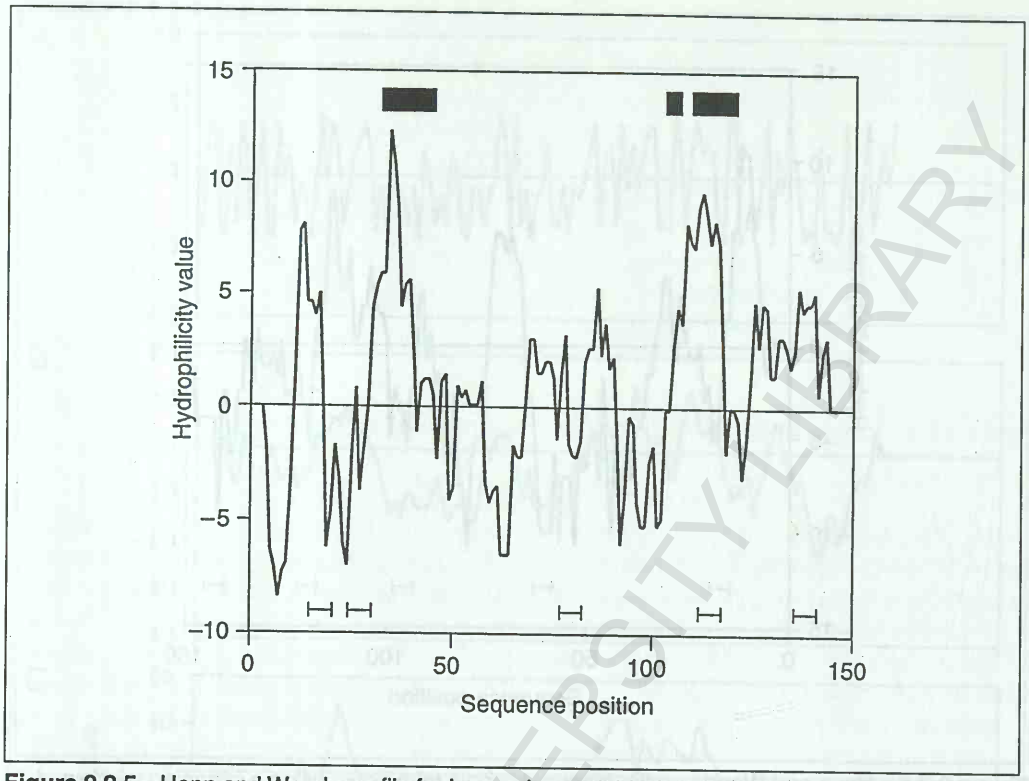


Figure 2.2.5 Hopp and Woods profile for human hemoglobin β-subunit using a scanning window size of 7 amino acids. Bars above the profile mark the locations of β-subunit surfaces in contact with α-subunits (Yoshioka and Atassi, 1986). Lines below the profile indicate the locations of antigenic sites.

above the profiles indicate the locations of β strands found in the crystal structure of this all-β protein (Finzel et al., 1989). Although there is a difference in scale magnitude between the Kyte and Doolittle and Eisenberg methods, both profiles show exceptionally good correlation of hydrophobicity (peak maxima) with the location of β strands. The correlation of hydrophobicity with β-sheet structural elements is a direct reflection of the fact that sheets are most often located in the hydrophobic interior of a protein; thus, they comprise residues most compatible with the hydrophobic environment. There is also good correlation between hydrophilicity (peak minima) and the location of turns in proteins. For interleukin 1β essentially all of the turns identified in the structure correspond to minima.

Figure 2.2.7 shows the hydropathy profiles for myoglobin, a protein composed solely of α helices. The larger α helices appear to correlate with peak maxima. Like β sheets, helices are frequently found in regions of hydrophobicity in the hydrophobic core of proteins. Because of the interspersed nature of hydrophilic and hydrophobic residues of amphipathic helices, however, hydropathy profiles are less useful in predicting the location of these helices. Moreover, as both helices and sheets result in peak maxima, the hydropathy profile is not sufficient to discriminate between them. Therefore, correlations of secondary structure with protein hydropathic character are most successful when used in conjunction with protein secondary structure prediction algorithms (see *UNIT 2.3*).

Predicting Transmembrane Regions in Proteins

Perhaps the most popular use of hydropathy profiles is predicting transmembrane regions in proteins. In addition to hydropathy scales, there are several membrane preference scales. Such scales have been developed independently from the hydropathy scales and are based on statistical distributions of each amino acid in membrane regions of proteins. The statistical scales have been previously compared to the hydropathy scales (Esposti et al., 1990); in general, they are equally efficient at predicting the location of the transmembrane and amphipathic helices.

Examining the structure of known membrane proteins has enabled the predictive rules

BU IAS CENTRAL UNIVERSITY LIBRARY

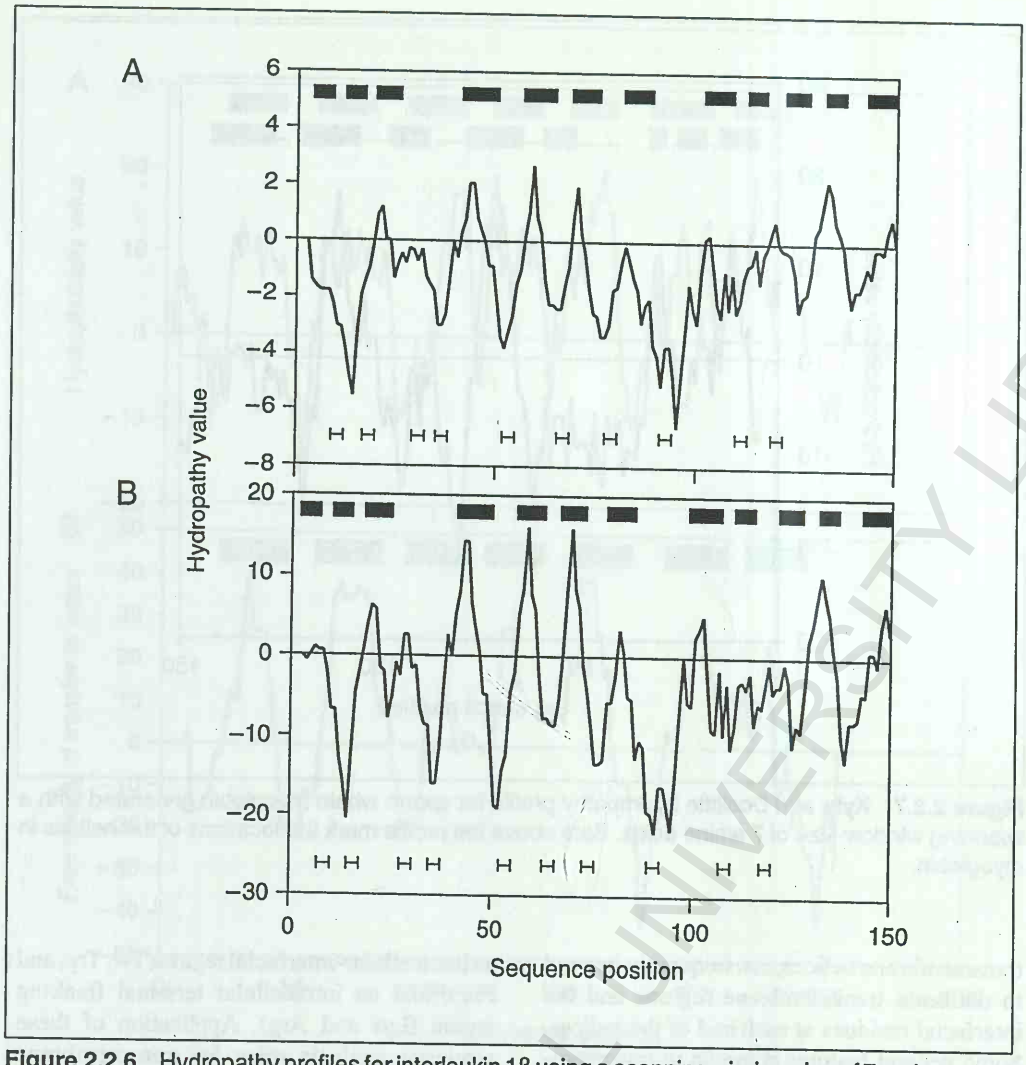


Figure 2.2.6 Hydropathy profiles for interleukin 1β using a scanning window size of 7 amino acids. Bars above the profile indicate locations of the β strands, and lines below the profile mark the locations of turns (Finzel et al., 1989). (**A**) Profile generated using the Eisenberg scale. (**B**) Profile generated using the Kyte and Doolittle scale.

for the various scales to be optimized. In addition, estimates of the hydrophobic thickness of the bilayer have provided guides to the minimum number of helical amino acids necessary to span the bilayer: the minimum number is estimated to be 20 (Engleman et al., 1986). This greatly facilitates the prediction of putative transmembrane helices by adding an additional criterion to the analysis of hydropathy profiles for transmembrane regions, namely, maxima in the profiles must span ~20 residues.

Figure 2.2.8A shows the Kyte and Doolittle profile for bacteriorhodopsin, a membrane protein whose structure is known (Henderson et al., 1990). For Kyte and Doolittle profiles, the best resolution is seen for window scans of 7 or 9 amino acids and threshold values of 0.7 × window size. For a window scan of 7, the threshold value is 4.9. There are seven major peaks with threshold values greater than 4.9 that span at least 20 amino acids (indicated by brackets above the profile in Fig. 2.2.8A). Comparison of the predicted transmembrane regions with the location of the actual helices identified in the structure (bars above the profile) indicates that the prediction is quite good. Similar agreement is seen when the method of Engleman et al. (1986) is used to predict the transmembrane regions for bacteriorhodopsin (Fig. 2.2.8B).

Although the location of the helices can be predicted reliably, neither the Kyte and Doolittle nor the GES method is able to predict accurately the boundaries of the transmembrane regions. Fortunately, visual inspection of the sequence, as well as general features of the

BOULDER CENTRAL UNIVERSITY LIBRARY

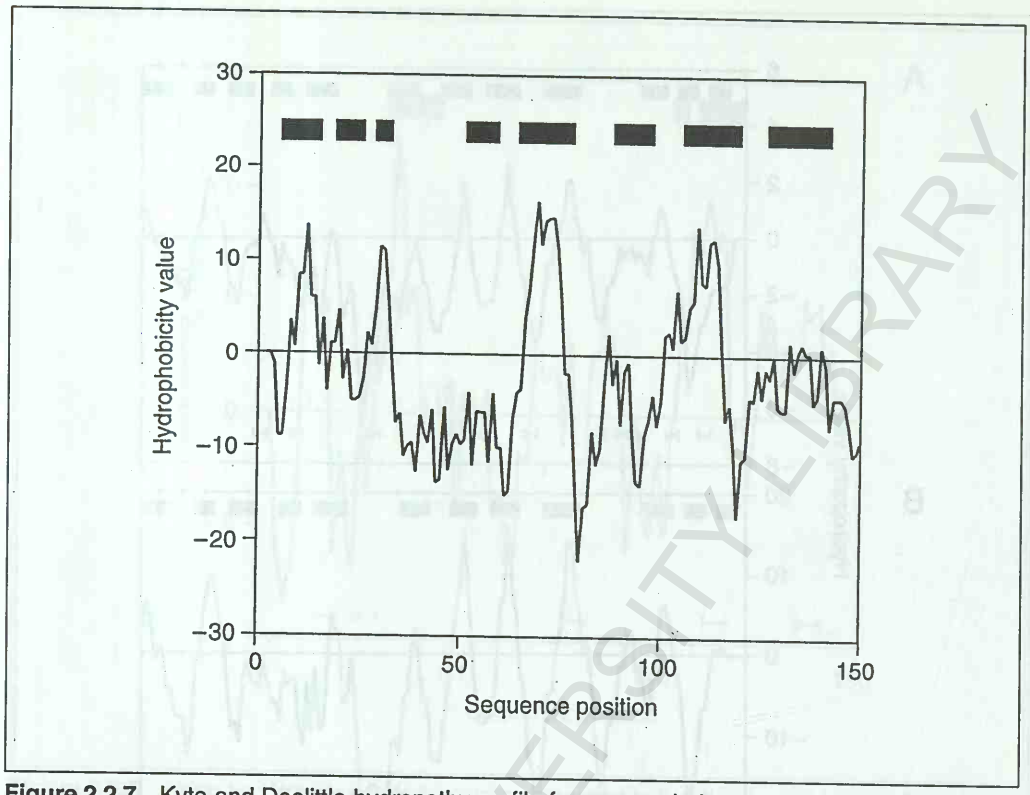


Figure 2.2.7 Kyte and Doolittle hydropathy profile for sperm whale myoglobin generated with a scanning window size of 7 amino acids. Bars above the profile mark the locations of the helices in myoglobin.

transmembrane helices, can frequently be used to delineate transmembrane regions and the interfacial residues at each end of the helices. Some general features common to transmembrane regions to consider are as follows. (1) Approximately 75% of the amino acids in transmembrane regions are hydrophobic (Leu, Ile, Val, Met, Phe, Trp, Tyr, Cys, Ala, Pro, and Gly). (2) Proline residues are found in ~20% of the transmembrane helices but in only 3% of globular proteins (Barlow and Thornton, 1988). (3) Aromatic residues are clustered near the interface of the transmembrane helix and bulk water.

Positional preferences for amino acids in transmembrane regions have been identified for single-spanning membrane proteins (Landolt-Marticorena et al., 1993) and can serve as a further guide for delineating the ends of transmembrane regions. It can be expected that these preferences will be similar for polytopic membrane proteins. The positional preferences define a motif with the following characteristics: an extracellular terminal flanking region (Asp, Ser and Pro); an extracellular-interfacial region (Trp); a transmembrane domain composed primarily of hydrophobic residues (Leu, Ile, Val, Met, Phe, Trp, Cys, Ala, Pro, and Gly); an intracellular-interfacial region (Tyr, Trp, and Phe); and an intracellular terminal flanking region (Lys and Arg). Application of these sequence analysis rules for transmembrane proteins greatly improves the results of using hydropathy profiles for predicting transmembrane regions.

CONCLUSIONS

1. Virtually all hydropathy scales are able to reveal structural information in an equivalent manner. Selection of the appropriate hydropathy scale is therefore a practical and/or personal decision.

2. Window scanning procedures should use an appropriate window size for the structural element being examined: 7 or 9 amino acids for predicting surface sites or secondary structure and 19 amino acids for predicting membrane-spanning regions.

3. Identification of interaction sites can be accomplished using peak minima (hydrophilicity).

4. Correlation of protein secondary structure with hydropathy profiles may be useful for protein structure prediction methodology (see *UNIT 2.3*).

RUTGERS UNIVERSITY LIBRARY

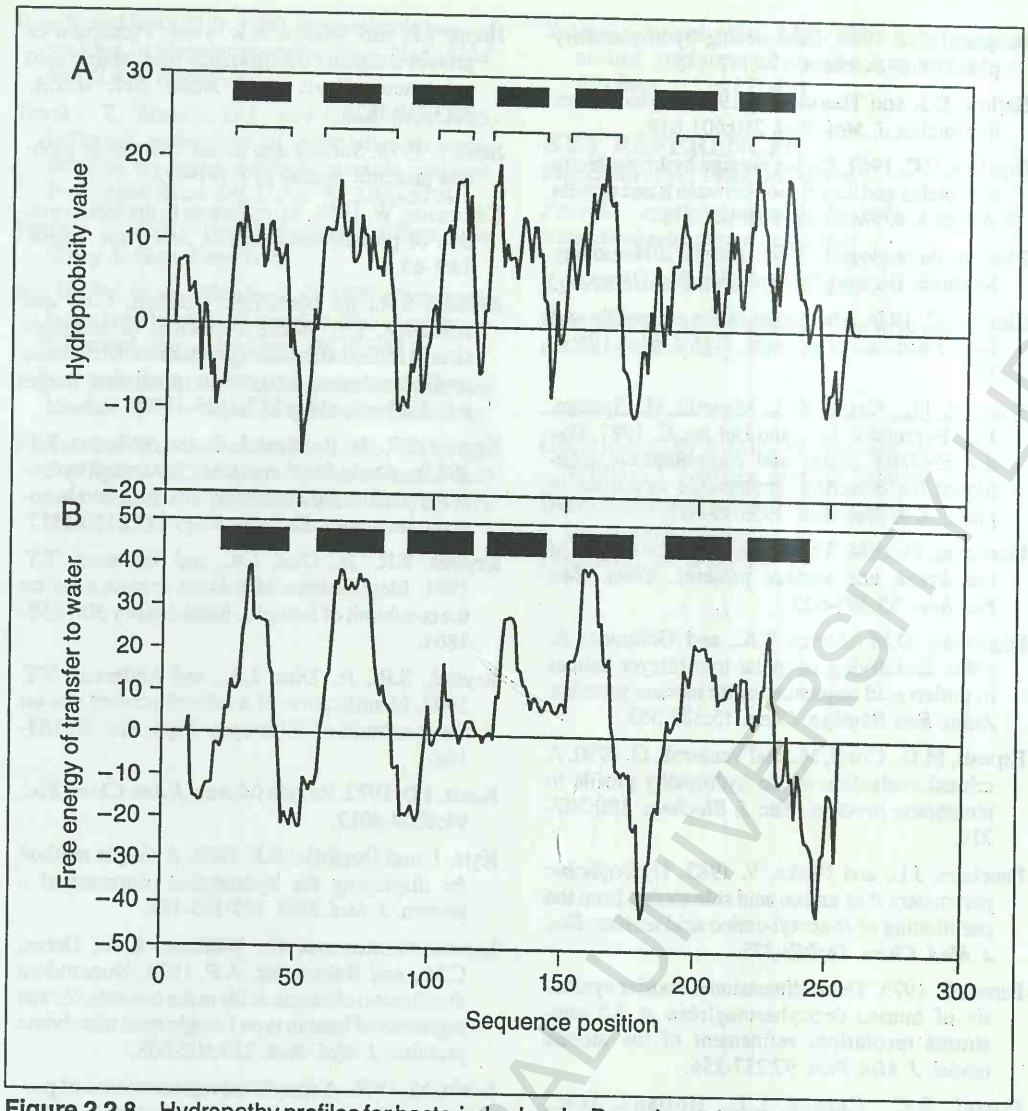


Figure 2.2.8 Hydropathy profiles for bacteriorhodopsin. Bars above the profiles mark the locations of the helices. (**A**) Kyte and Doolittle profile generated using a scanning window size of 7 amino acids. Brackets indicate the helices predicted on the basis of the profile alone. (**B**) GES profile generated using a scanning window size of 19 amino acids.

5. Prediction of transmembrane regions of integral membrane proteins can be made using a variety of scales and protocols; analysis of the hydropathy profiles should be augmented with the applicable sequence analysis rules.

ACCESSIBILITY OF SOFTWARE

Most commercial personal computer or computational packages (Genetics Computer Group, 1994; Biosym Technologies, 1991; Tripos Associates, 1993) contain algorithms for hydropathy profile generation along with other secondary structure prediction tools (see Table 2.3.6). In addition, the authors of each scale or protocol have published analysis programs (Table 2.2.1), of which many are available from public domain sources on the Internet. Algorithms can also easily be encoded into programming languages or spreadsheets and interfaced with numerous graphical software packages for plotting the hydropathy profiles (if graphics are not included in the particular prediction program). These are also discussed in *UNIT 2.1*.

LITERATURE CITED

Anfinsen, C.B., Haber, E., Sela, M., and White, F.H., Jr. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* 47:1309-1314.

Argos, P., Rao, J.K.M., and Hargrave, P.A. 1982. Structural prediction of membrane-bound proteins. *Eur. J. Biochem.* 128:565-575.

Atassi, M.Z. 1984. Antigenic structure of proteins. *Eur. J. Biochem.* 145:1-20.

CENTRAL UNIVERSITY LIBRARY

Banghan, J.A. 1988. Data-sieving hydrophobicity plots. *Anal. Biochem.* 174:142-145.

Barlow, D.J. and Thornton, J. 1988. Alpha helices in proteins. *J. Mol. Biol.* 210:601-619.

Bigelow, C.C. 1967. On the average hydrophobicity of proteins and the relation between it and protein structure. *J. Theor. Biol.* 16:187-211.

Biosym Technologies. 1991. Insight II/Homology Modules. Biosym Technologies, San Diego.

Chothia, C. 1976. The nature of the accessible and buried surfaces of proteins. *J. Mol. Biol.* 105:1-14.

Cornette, J.L., Cease, K.B., Margalit, H., Spouge, J.L., Berzofsky, J.A., and DeLisa, C. 1987. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J. Mol. Biol.* 195:659-685.

Eisenberg, D. 1984. Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* 53:595-623.

Engleman, D.M., Steitz, T.A., and Goldman, A. 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Chem.* 15:321-353.

Esposti, M.D., Crimi, M., and Venturoli, G. 1990. A critical evaluation of the hydropathy profile in membrane proteins. *Eur. J. Biochem.* 190:207-219.

Fauchere, J.L. and Pliska, V. 1983. Hydrophobic parameters *P* of amino acid side chains from the partitioning of *N*-acetyl-amino acid amides. *Eur. J. Med. Chem.* 18:369-375.

Fermi, G. 1975. Three-dimensional Fourier synthesis of human deoxyhaemoglobin at 2.5 angstroms resolution, refinement of the atomic model. *J. Mol. Biol.* 97:237-256.

Finzel, B.C., Clancy, L.L., Holland, D.R., Muchmore, S.W., Watenpaugh, K.D., and Einspahr, H.M. 1989. Crystal structure of recombinant human interleukin-1β at 2.0 angstroms resolution. *J. Mol. Biol.* 209:779-791.

Fraga, S. 1982. Theoretical prediction of protein antigenic determinants from amino acid sequences. *Can. J. Chem.* 60:2606-2610.

Genetics Computer Group. 1994. GCG Program Manual for the Wisconsin Package. Genetics Computer Group, Inc., Madison, Wis.

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., and Downing, K.H. 1990. Model for the structure of bacteriorhodopsin based on high resolution electron cryomicroscopy. *J. Mol. Biol.* 213:899-929.

Hopp, T.P. 1984. Protein antigen conformation: Folding patterns and predictive algorithms; selection of antigenic and immunogenic peptides. *Ann. Sclavo* 2:47-60.

Hopp, T.P. 1989. Use of hydrophilicity plotting procedures to identify protein antigenic segments and other interaction sites. *Methods Enzymol.* 178:571-585.

Hopp, T.P. and Woods, K.R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* 78:3824-3828.

Janin, J. 1979. Surface and inside volumes in globular proteins. *Nature* 277:491-492.

Kauzmann, W. 1959. Some factors in the interpretation of protein denaturation. *Adv. Prot. Chem.* 14:1-63.

Krystek, S.R., Jr., Dias, J.A., Reichert, L.E., and Andersen, T.T. 1985a. Prediction of antigenic sites in follicle stimulating hormones: Difference profiles enhance antigenicity prediction methods. *Endocrinology* 117:1125-1130.

Krystek, S.R., Jr., Reichert, L.E., and Andersen, T.T. 1985b. Analysis of computer generated hydropathy profiles for human glycoprotein and lactogenic hormones. *Endocrinology* 117:1110-1117.

Krystek, S.R., Jr., Dias, J.A., and Andersen, T.T. 1991. Identification of subunit contact sites on the α-subunit of lutropin. *Biochemistry* 30:1858-1864.

Krystek, S.R., Jr., Dias, J.A., and Andersen, T.T. 1992. Identification of a subunit contact site on the α-subunit of follitropin. *Pept. Res.* 25:165-168.

Kuntz, I.D. 1972. Protein folding. *J. Am. Chem. Soc.* 94:4009-4012.

Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

Landolt-Marticorena, C., Williams, K.A., Deber, C.M., and Reithmeier, A.F. 1993. Nonrandom distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. *J. Mol. Biol.* 229:602-608.

Levitt, M. 1976. A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* 104:59-107.

Nozaki, Y. and Tanford, C. 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. *J. Biol. Chem.* 246:2211-2217.

Parker, J.M.R., Guo, D., and Hodges, R.S. 1986. New hydrophobicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 25:5425-5432.

Rao, J.K.M. and Argos, P. 1986. A conformational preference parameter to predict helices in integral membrane proteins. *Biochim. Biophys. Acta* 869:197-214.

Rose, G.D. and Dworkin, J.E. (1989). The hydrophobicity profile. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 625-633. Plenum, New York.

Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H., and Zehfus, M.H. 1985. Hydrophobicity of amino acid residues in globular proteins. *Science* 229:834-838.

Roy, S. and Rose, G.D. 1980. Hydrophobic basis of packing in globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* 8:4643-4647.

Tanaka, T., Slamon, D.J., and Cline, M.J. 1985. Efficient generation of antibodies to oncoproteins by using synthetic peptide antigens. *Proc. Natl. Acad. Sci. U.S.A.* 82:3400-3404.

Tanford, C. 1980. The Hydrophobic Effect. John Wiley & Sons, New York.

von Heijne, G. and Blomberg, C. 1979. Transmembrane translocation of proteins. The direct transfer model. *Eur. J. Biochem.* 97:175-181.

Tripos Associates. 1993. Biopolymer/Composer Module. Tripos Associates, Inc., St. Louis, Mo.

Wolfenden, R., Andersson, L., Cullis, P.M., and Southgate, C.C. 1981. Affinities of amino acid side chains for solvent water. *Biochemistry* 20:849-855.

Yoshioka, N. and Atassi, M.Z. 1986. Subunit interacting surfaces of human haemoglobin. *Biochem. J.* 234:457-461.

KEY REFERENCES

Engleman et al., 1986. See above.

Provides applications of hydropathy profiles to transmembrane region prediction.

Kyte and Doolittle, 1982. See above.

Describes the fundamental development and application of hydropathy profiles.

Contributed by Stanley R. Krystek, Jr., William J. Metzler, and Jiri Novotny
Bristol-Myers Squibb Pharmaceutical Research Institute
Princeton, New Jersey

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Protein Secondary Structure Prediction

It is widely accepted that knowledge of the structure of a protein can provide important information regarding its function and mechanism of action. However, determining the three-dimensional structure of a protein, whether by X-ray crystallography (Wyckoff et al., 1985; Blundell and Johnson, 1976) or nuclear magnetic resonance (NMR) spectroscopy (Oppenheimer and James, 1989; Wuthrich, 1986), remains a formidable task, as it requires relatively large amounts of pure protein (generally greater than milligram quantities). Thus, the structures of many proteins will remain out of reach. At the same time, continual advances in molecular biology provide protein sequence information (primary structures) at a pace that far exceeds the speed with which higher-order protein structures can be determined, thus making the need for methods to derive structural information even more pressing. This unit describes procedures developed for predicting protein structure from the amino acid sequence.

On the basis of function and location in the cell, proteins can be loosely divided into three groups: globular, structural, and membrane. To date, the majority of protein structural data has been derived from studies of globular (soluble) proteins because of the ease of chemical handling of globular proteins. Thus, a wide disparity has developed in the number of three-dimensional structures that have been experimentally determined in the three protein groups. As a result, current secondary structure prediction schemes apply almost solely to globular proteins.

In the native structures of globular proteins, the hydrophobic residues are generally buried in the interior and the hydrophilic residues are exposed to the surrounding solvent (Kauzmann, 1959; Fisher, 1964; Dill, 1990). Disulfide bonds are often present and stabilize the structure. Sections of amino acid sequence rich in hydrophobic residues can be correlated with secondary structural elements such as α helices and β strands (Roy and Rose, 1980). Those rich in polar residues correspond to surface loops and β turns (Kuntz, 1972; *UNIT 2.2*). These secondary structural elements (Pauling and Corey, 1951; Pauling et al., 1951) are characterized by hydrogen bonding between the amino and carboxyl groups of the polypeptide backbone (Schulz and Schirmer, 1979).

Several reviews of protein structure exist, and readers are referred to books by Schulz and Schirmer (1979) and Branden and Tooze (1991) for thorough discussions of protein structure. Protein secondary structure elements associate to form compact domains. Levitt and Chothia (1976) and Richardson (1981) have classified protein domains according to the types of arrangements of secondary structures that form them. Illustrations for the four classes, referred to as all-α, all-β, α/β, and α+β, are available (Richardson, 1981; Lesk, 1991). This unit focuses on examining secondary structure predictions using specific examples from different protein classes; methods of predicting folding class are also discussed.

This unit is divided into four sections. The first section is an overview and brief history of structure prediction schemes. The second section describes four distinct prediction schemes, with emphasis on their differences. In the third part each prediction scheme is used to evaluate three proteins that have different folding patterns. The final section is a comparison of the prediction results and suggestions for secondary structure prediction.

OVERVIEW OF PREDICTION SCHEMES

Early Predictive Schemes

Two of the earliest secondary structure prediction methods were developed by Guzzo (1965) and Prothero (1966). By analyzing a database of protein structures and corresponding sequences, they generated a set of rules for predicting structure. Because of the limited number of structures available at the time, their prediction methods were restricted to helices, the most abundant conformation in the structures available.

Soon after, Schiffer and Edmunson (1967) developed the helical wheel method, a pictorial tool that is still widely used. The helical wheel is a circular plot that represents the view one obtains when looking down the helix axis (Fig. 2.3.1). The "spokes" of the wheel are precisely positioned on the basis of the Corey-Pauling α-helical geometry, having 3.6 residues per turn, which results in an angular shift of two neighboring residues of 100° (Pauling et al., 1951). A helical wheel is constructed by placing successive amino acids at the spokes of the wheel.

Helical wheels have been particularly valuable for identifying proteins and peptides that contain amphipathic helices. In an amphipathic

Copyright © 1995 by John Wiley & Sons, Inc.

BCM CENTRAL LIBRARY

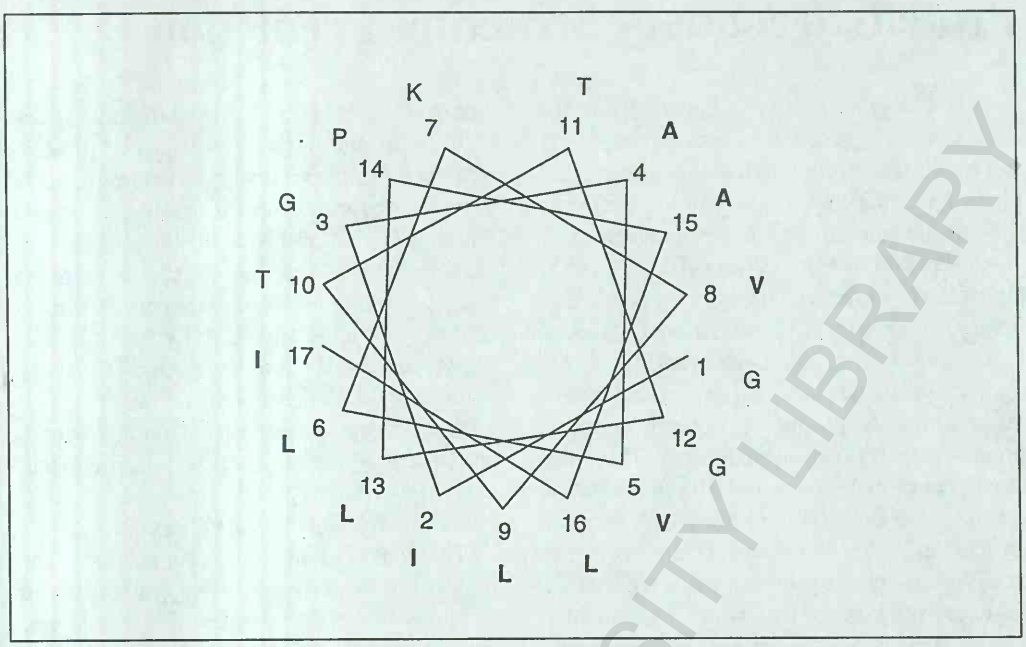


Figure 2.3.1 Helical wheel for a portion of the amphipathic helix of melittin. Residues in bold type form a hydrophobic side. Note that this is not a perfect amphipathic helix as the polar face includes Ala-4, Ala-15, and Val-8. This helix also contains several glycine residues, which are not commonly found in helices (Table 2.3.1).

helix, hydrophobic residues will cluster to one side of the wheel and polar residues will be found on the opposite side. Figure 2.3.1 shows a helical wheel for the amphipathic helix of melittin. Helical wheels have often been used for analyzing helical dimerization proteins such as leucine zippers (Krystek et al., 1991; O'Shea et al., 1992), for designing helical multimers (Harbury et al., 1993; Kamtekar et al., 1993; Lovejoy et al., 1993), and for predicting and analyzing membrane proteins which contain helices that span the lipid bilayer (Baldwin, 1993).

Other early predictive schemes include those developed by Dunhill (1968) and Lim (1974). Dunhill's "helical net" is a pictorial method that is similar to the helical wheel and is still used in protein design and helix prediction (DeGrado, 1988; Goodman and Kim, 1991). Lim's method relies on an extensive set of rules derived from the physicochemical properties of amino acids in conjunction with the analysis of a structural database. The large number of detailed rules, however, makes the Lim method unwieldy. When several methods were compared in a blind test organized by Schulz et al. (1974), the methods of Ptitsyn and Finkelstein (1970) and Chou and Fasman (1974a,b) performed better than others (Burgess et al., 1971; Lewis et al., 1971; Robson and Pain, 1971; Nagano, 1973).

Statistical Prediction Schemes

The statistical methods developed by Chou and Fasman (1974a,b) and by Garnier et al. (1978) are the most prominent of the class of predictive schemes based on statistical amino acid preference rules. These secondary structure prediction schemes tacitly assume that only local sequence interactions influence the formation of local structure. This is almost certainly an oversimplification. Significant interactions occur between protein secondary structural elements (tertiary structure); these interactions may involve hydrophobic and/or electrostatic interactions between structural units. Therefore, the success and accuracy of prediction schemes that rely on empirical algorithms are necessarily limited and are unlikely to exceed a level of 70% to 75%. Reports of secondary structure prediction accuracy levels of ~70%, however, are not uncommon using these methods (Fasman, 1985).

Other Predictive Schemes

Two additional methodologies for structure prediction not based on statistics of amino acid occurrence have evolved. Of the neural network methods (Quinn and Sejnowski, 1988; Rost and Sander, 1993a,b; Kneller et al., 1990), that of Rost and Sander claims a level of predictive accuracy exceeding 70% for all proteins tested and is publicly available through the

RCU IASTATE UNIVERSITY LIBRARY

Internet computer network (for details, see discussion of Applying the PHD Scheme and Table 2.3.6 therein). The method of Stultz et al. (1993) is based on state-space modeling with a set of predefined structural class models.

Selecting Prediction Tools

Secondary structure prediction proceeds as follows. First, it should be confirmed that the three-dimensional structure has not yet been determined. This can be done by searching known X-ray and NMR structures in the Protein Data Bank (Brookhaven, 1994) for an amino acid sequence identical to that of interest. Next, identification of proteins with significantly similar sequences whose structures are already known should be attempted. Numerous software packages are available for performing similarity searches (see *UNIT 2.1*). If proteins of similar sequence (>20% identity) are found, the Protein Data Bank should again be searched to determine whether the structures of any of those proteins are known. If so, then comparative homology modeling (Greer, 1990; Bajorath et al., 1993) can be used to develop a three-dimensional model of the desired sequence. Only if there are no proteins in the structural database of sufficient similarity should protein secondary prediction schemes be used. It should be pointed out, however, that proteins may have the same fold without a detectable sequence similarity (Kabsch and Sander, 1984, 1985). In the future, sequence threading through known three-dimensional structures will be helpful (Jones et al., 1992; Sippl and Weitckus, 1992).

It is recommended that protein secondary structure prediction be performed in conjunction with other sequence analysis approaches, such as hydrophobicity profiling, flexibility analysis, surface exposure experiments, charge distribution mapping, and motif searches (*UNIT 2.2*; Fasman, 1989b; Genetics Computer Group, 1994). For investigators using methods to predict protein secondary structure, at least two questions should be addressed: (1) which predictive scheme(s) should be used, and (2) what are the advantages of one method over others? This unit attempts to find answers to these questions by studying three proteins in detail: human hemoglobin β-subunit, human interleukin 1β (IL-1β), and human profilin I (see Application of Secondary Prediction Methods and Fig 2.3.2).

METHODS FOR SECONDARY STRUCTURE PREDICTION

The Chou and Fasman Method

The most popular secondary structure prediction scheme is the statistical method of Chou and Fasman (1974a). Using a database of 15 protein structures determined by X-ray crystallography, Chou and Fasman tabulated the number of occurrences of a given amino acid in α helix, β strand, or coil. From the tabulation, a set of conformational preference parameters for each amino acid type was calculated. The preference parameters were based on three functions: (1) the relative frequency of a specific amino acid type within each protein, (2) the occurrence of that amino acid in a given type of secondary structure, and (3) the fraction of amino acid residues occurring in each type of secondary structure. With the conformational preference parameters in hand, the authors then derived a set of empirical rules for predicting protein secondary structure (Chou and Fasman, 1974b).

Chou and Fasman later published an extension of their analysis using a larger set of X-ray structures, totaling 29 proteins and over 4700 amino acids. In addition to refining the values for α helices and β strands, they included a conformational preference parameter for β turns as well (Chou and Fasman, 1977, 1978a,b). For β turns, a positional preference for residues was identified; for example, proline was found to occur more frequently in position 2 of the turn than in positions 1, 3, or 4.

A further refined set of conformational preference parameters for helical and sheet residues was later derived from an expanded data set of 64 protein structures (Chou, 1989). These parameters are similar to those based on the 29 protein structures (differences are noted in Chou, 1989). An interesting addition to secondary structure prediction using the conformational preference parameters is the prediction of protein class. Chou presented a computerized algorithm that assigns a structural class to proteins on the basis of amino acid composition; however, these predictions have not been widely used.

In the initial studies, Chou and Fasman used the conformational preference parameters to describe amino acids as being formers, breakers, or indifferent for helical and strand regions (Chou and Fasman, 1974a). Their analysis showed that residues with the highest prefer-

BOSTON COLLEGE LIBRARY

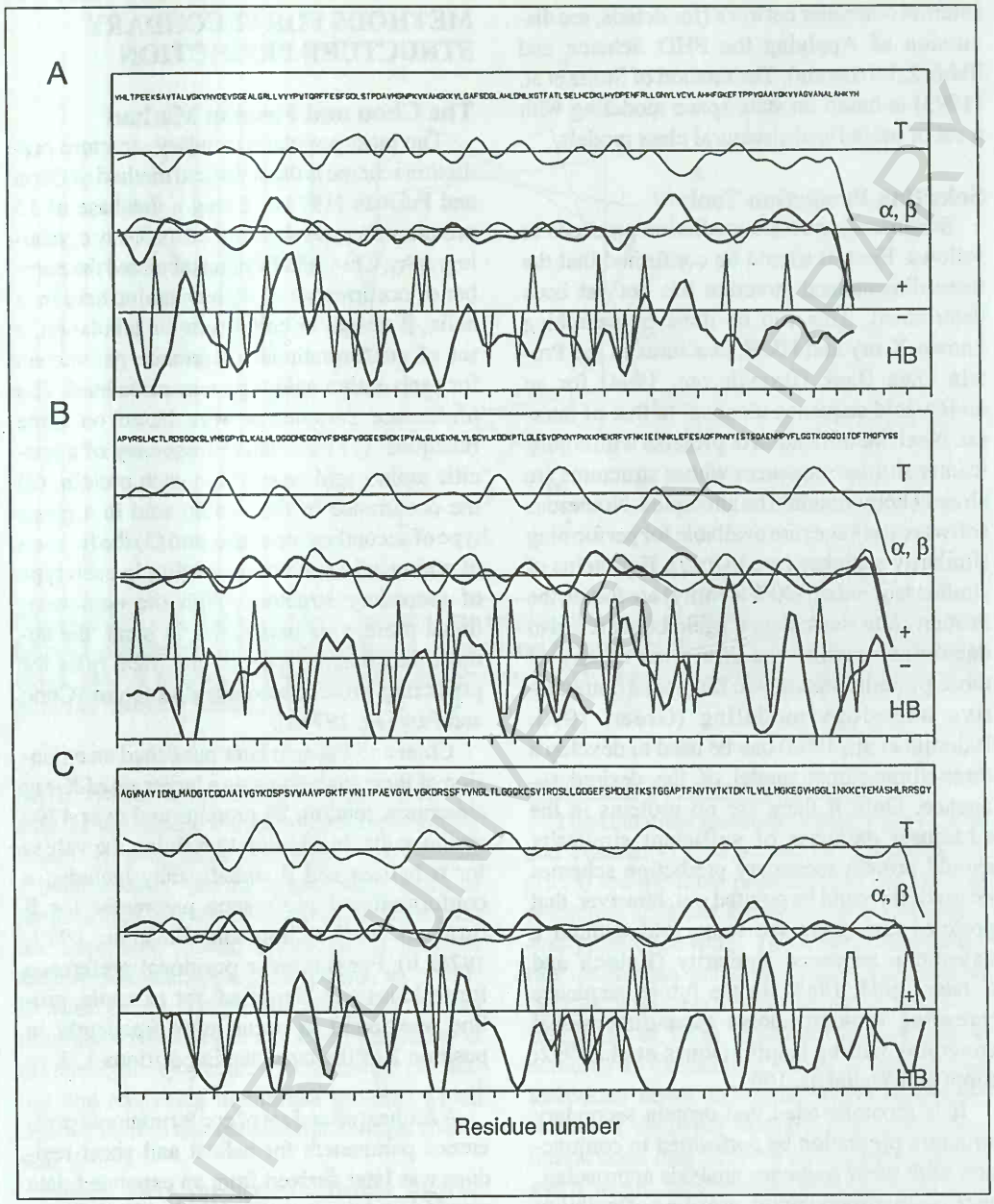


Figure 2.3.2 Secondary structure prediction using the method of Chou and Fasman as implemented by Novotny and Auffray (1984). Predicted secondary structure for (**A**) human hemoglobin β-subunit, (**B**) interleukin 1β, and (**C**) human profilin I. Below the amino acid sequence are given the turn prediction (T), the α helix (thin line) and β strand prediction (thick line), the location of positive (upward spikes) or negative charges (downward spikes) as derived specifically from amino acid side chains, and a hydrophobicity profile (HB).

ence for helices resided near the helix center, whereas residues with low helix preference could be found clustered at the helix termini. For helix and strand, residues were independently classified as strong formers, formers, weak formers, indifferent, breakers, and strong breakers. Table 2.3.1 contains a list of the conformational preference parameters and refined conformational assignments derived from the analysis of 64 proteins (Chou, 1989). These values are combined with the turn preferences shown in Table 2.3.2, or values from the 29-protein study (Chou and Fasman, 1978a,b; Prevelige and Fasman, 1989), to carry out sequence prediction as described below.

An important contribution to helix prediction strategies is the recognition that helix start and stop signals in the amino acid sequence grammar are quite distinct (Presta and Rose, 1988; Richardson and Richardson, 1988). The

BCU IASRI CENTRAL UNIVERSITY LIBRARY

Table 2.3.1 Conformational Preference and Assignments for α Helix and β Strand[a]

α helix			β strand		
Residue	Pα	Assignment[b]	Residue	Pβ	Assignment[c]
Glu	1.44	Hα	Val	1.64	Hβ
Ala	1.39	Hα	Ile	1.57	Hβ
Met	1.32	Hα	Thr	1.33	hβ
Leu	1.30	Hα	Tyr	1.31	hβ
Lys	1.21	hα	Trp	1.24	hβ
His	1.12	hα	Phe	1.23	hβ
Gln	1.12	hα	Leu	1.17	hβ
Phe	1.11	hα	Cys	1.07	hβ
Asp	1.06	hα	Met	1.01	Iβ
Trp	1.03	Iα	Gln	1.00	Iβ
Arg	1.00	Iα	Ser	0.94	iβ
Ile	0.99	iα	Arg	0.94	iβ
Val	0.97	iα	Gly	0.87	iβ
Cys	0.95	iα	His	0.83	iβ
Thr	0.78	iα	Ala	0.79	iβ
Asn	0.78	iα	Lys	0.73	bβ
Tyr	0.73	bα	Asp	0.66	bβ
Ser	0.72	bα	Asn	0.66	bβ
Gly	0.63	Bα	Pro	0.62	Bβ
Pro	0.55	Bα	Glu	0.51	Bβ

[a]Values taken from Chou, 1989.

[b]Hα, strong helix former; hα, helix former; Iα, weak helix former; iα, indifferent; bα, helix breaker; Bα, strong helix breaker.

[c]Hβ, strong β-strand former; hβ, β-strand former; Iβ, weak β-strand former; iβ, indifferent; bβ, β-strand breaker; Bβ, strong β-strand breaker

α-helical structures in proteins are commonly stabilized by reciprocal hydrogen bond formation between the unpaired main-chain amino and carboxyl groups near the helix termini and the side chains of amino acids flanking the termini (Harper and Rose, 1993). The helix-terminal residues are referred to as the N-cap and C-cap. Because of the unique geometry of proline, this amino acid is frequently found at the N-cap + 1 position and has been referred to as being an α-helix initiator (Richardson and Richardson, 1988). Residues that typically occupy the N-cap and C-cap positions of helices are listed in Table 2.3.3.

The Garnier, Osguthorpe, and Robson Method

The Garnier, Osguthorpe, and Robson (GOR) method (Garnier et al., 1978; Garnier and Robson, 1989) is a widely available secondary structure prediction method that has been incorporated into several computer packages, such as that from Genetics Computer Group (GCG). Although the GOR method incorporates essentially all the elements of the Chou-Fasman statistics, it also uses information theory to reach the final decision. A detailed description of the GOR method can be found elsewhere (Garnier and Robson, 1989).

In principle, the GOR method assigns the conformation of a given residue by considering the contribution of residues relatively distant to it in the sequence. On the basis of studies showing the interdependence of residue conformations up to eight amino acids away, the GOR method assigns single-residue information values to each amino acid along the polypeptide chain. The predicted conformational state for an amino acid is the state with the highest positive information value. An extension of this procedure, the GORIII method, has been developed to use pair information instead of directional information. The GORIII method extracts both single-residue information values

BUCHANAN CENTRAL UNIVERSITY LIBRARY

Table 2.3.2 Conformational Preference and Assignments for β Turns[a]

Residue	Preference	Positional preferences for β turns[b]							
		f_i		f_{i+1}		f_{i+2}		f_{i+3}	
Asn	1.56	Asn	0.161	Pro	0.301	Asn	0.191	Trp	0.167
Gly	1.56	Cys	0.149	Ser	0.139	Gly	0.190	Gly	0.152
Pro	1.52	Asp	0.147	Lys	0.115	Asp	0.179	Cys	0.128
Asp	1.46	His	0.140	Asp	0.110	Ser	0.125	Tyr	0.125
Ser	1.43	Ser	0.120	Thr	0.108	Cys	0.117	Ser	0.106
Cys	1.19	Pro	0.102	Arg	0.106	Tyr	0.114	Gln	0.098
Tyr	1.14	Gly	0.102	Gln	0.098	Arg	0.099	Lys	0.095
Lys	1.01	Thr	0.086	Gly	0.085	His	0.093	Asn	0.091
Gln	0.98	Tyr	0.082	Asn	0.083	Glu	0.077	Arg	0.085
Thr	0.96	Trp	0.077	Met	0.082	Lys	0.072	Asp	0.081
Trp	0.96	Gln	0.074	Ala	0.076	Thr	0.065	Thr	0.079
Arg	0.95	Arg	0.070	Tyr	0.065	Phe	0.065	Leu	0.070
His	0.95	Met	0.068	Glu	0.060	Trp	0.064	Pro	0.068
Glu	0.74	Val	0.062	Cys	0.053	Gln	0.037	Phe	0.065
Ala	0.66	Leu	0.061	Val	0.048	Leu	0.036	Glu	0.064
Met	0.60	Ala	0.060	His	0.047	Ala	0.035	Ala	0.058
Phe	0.60	Phe	0.059	Phe	0.041	Pro	0.034	Ile	0.056
Leu	0.59	Glu	0.056	Ile	0.034	Val	0.028	Met	0.055
Val	0.50	Lys	0.055	Leu	0.025	Met	0.014	His	0.054
Ile	0.47	Ile	0.043	Trp	0.013	Ile	0.013	Val	0.053

[a]Values taken from Chou and Fasman (1977, 1978a,b).

[b]The positional preference f_i is defined as the relative occurence of a specific amino acid found in position i of a β turn.

Table 2.3.3 Amino Acid Positional Preferences in α Helices[a]

Effect	Residues[b]
Amino-terminal stabilizing (N-cap)	Gly Ser, Thr Asp, Glu Asn, Gln His
Carboxyl-terminal stabilizing (C-cap)	Gly Asn His(+), Lys, Arg
Amino-terminal destabilizing	His(+), Lys, Arg Val, Leu, Ile, Met
Carboxyl-terminal destabilizing	Ser Asp Val, Leu, Ile, Met

[a]Amino acid preferences were compiled from Fersht and Serrone (1993), Presta and Rose (1988), and Richardson and Richardson (1988).

[b]Residues listed in order of increasing magnitude of the effect.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

and pair information values for each possible pair and for amino acids within eight amino acids of a central residue. The predictive accuracy of the GOR methods is similar to that of the Chou and Fasman method (Garnier et al., 1978). Thus, the GOR methods may serve as a check of the results obtained with the Chou and Fasman method.

The EMBL Profile Neural Network Method

An important feature of the EMBL Profile Neural Network (PHD) prediction scheme is the use of multiple sequence alignments (Rost and Sander, 1993a,b). Previous studies have shown that homologous proteins (generally with at least 20% to 30% sequence identity but often less) frequently have identical three-dimensional folds (tertiary structure). Thus, multiple sequence alignments are grouped by structural family and used as input. By considering many aligned sequences, the algorithm makes use of more information about structure than is available from just the individual protein sequence for which the secondary structure is to be predicted.

To determine which protein sequences are to be used in the prediction, the PHD method uses a weighted dynamic programming method, MaxHom (Sander and Schneider, 1991), in searching a database of sequences. A family profile for the input sequence is generated and the amino acid frequencies at each alignment position are calculated for use in the prediction. Results have shown that the multiple sequence alignments improve secondary structure prediction >6% compared to single-sequence predictions (Rost and Sander, 1993a,b).

The PHD method uses a **reference network** or **neural network** that has been trained and tested on a database of nonhomologous protein structures. In a typical database there is an uneven distribution of secondary structure types, such that there are many more examples of turns than of helices or sheets (Rost and Sander, 1993a,b). Because of the uneven distribution, empirically based predictions of loops are more accurate than those of helices, which in turn are more accurate than those of sheets. The neural network approach of the PHD method eliminates this bias: Rost and Sander developed a network that is trained with each type of secondary structure in equal proportion, rather than the proportions present in any database of protein structures. This results in a more balanced prediction.

The network system for structure prediction consists of three layers. In the first layer, secondary structure is predicted on the basis of the multiple aligned sequences. The results are passed to the second layer, which collectively considers secondary structure elements that are adjacent in the protein sequence, thereby allowing nearest neighbor structure to influence the structure of a particular region of the protein. The third layer of the network averages the results of the independently trained networks. The output from the third layer is determined by a "winner take all" prediction of the secondary structure at a given sequence position (Rost and Sander, 1993a,b).

The PSA Method

The Protein Sequence Analysis (PSA) method (Stultz et al., 1993) computes the probability that a given amino acid belongs to a particular secondary structure element when there are no homologous protein sequences or structural data. To do this, mathematical models were developed for 15 single-domain protein **superclasses** or **macroclasses**. These superclasses, patterned after the protein class categorization of Richardson (1981), are groups of proteins with similar secondary and tertiary structures. The mathematical models represent constraints on the patterns of secondary structure elements for a given protein class.

To perform structure prediction, the PSA method uses the amino acid sequence to calculate the probability that a given residue is contained in each of the modeled secondary structure elements. Also calculated is the probability that the submitted sequence is a member of one of the predefined superclasses. Thus, in addition to evaluating the secondary structure prediction, the PSA method provides the bonus of the potential tertiary structural information contained in the predicted structural superclass.

The Helical Wheel

The helical wheel is a plot of the amino acid residues around a potentially helical segment (see Fig. 2.3.1 and discussion in Early Predictive Schemes). The method was developed to find helices with a hydrophobic face buried away from a polar solvent, with the graphical representation showing the clustering of polar and/or nonpolar residues toward one face of a helix. In its original use the helical wheel facilitated the identification of potential helical segments in protein sequences (Schiffer and Edmundson, 1967), but its applicability has since been expanded to include designing proteins (e.g., leu-

BOSTON UNIVERSITY LIBRARY

cine zipper proteins; Harbury et al., 1993; Kamtekar et al., 1993; Lovejoy et al., 1993) and studying transmembrane proteins (e.g., G-protein-coupled receptors; Baldwin, 1993).

APPLICATION OF SECONDARY STRUCTURE PREDICTION METHODS

Applying the Chou and Fasman Method

The wide use of the Chou and Fasman method is due in part to the simplicity of application, ease of understanding, and potential interactive nature of the analysis. It also ranks favorably in accuracy of prediction relative to other methods. Application of the method entails the following steps.

1. Conformational preference parameters for each type of secondary structure are locally smoothed by averaging techniques (see *UNIT 2.2* for an example of the scanning window averaging method) in order to calculate α-helix, β-sheet, and turn probabilities along the polypeptide backbone.

2. The α-helix, β-strand, and turn probabilities are plotted versus sequence position for interpretation. Basically, the conformation is assigned according to the highest curve (Fig. 2.3.2, Fig. 2.3.3, and Fig. 2.3.4).

3. On the basis of conformational preference parameters, amino acids are assigned as formers, breakers, or indifferent for α-helical and β-sheet regions (Fig. 2.3.3A).

4. Program options usually include full predictions that are based on the original Chou and Fasman rules (Table 2.3.4). Predictions can also be made by the user in an interactive environment by combining the Chou and Fasman rules with sequence analysis using other empirical methods such as hydrophobicity profiling, flexibility prediction, charged residue mapping, and surface probability calculation (Fig. 2.3.3 and Fig. 2.3.4; see also Comparison of Prediction Schemes). We find that an interactive approach allows for the best interpretation of the prediction results (Fasman, 1989a; Schulz, 1988). Several other correlations can be used in this analysis and are listed in Table 2.3.5. Many programs can also generate sketches of predicted peptide structures (Fig. 2.3.5).

Applying the GOR Scheme

The GOR scheme is coded into many commercial and academic protein structure analysis packages. This method provides estimated probabilities that a given residue is in a particular conformational state, and it predicts the conformation of an amino acid without any input from the user other than the sequence of interest. There are two outputs from the GOR method. The first is a plot of the turns, α helices, and β strands as predicted from the method. The second is a summary of the full protein structure prediction similar to that produced by the Chou and Fasman method. Unlike the Chou and Fasman method, assignment of each amino acid to a structural element is made automatically. Because many prediction schemes can be printed in the same output format, it is possible to print results of Chou and Fasman, GOR, and other structural predictions together for a more interactive analysis of the sequence (Fig. 2.3.4). Use of a combination of structural prediction schemes and the secondary structure correlations presented in Table 2.3.5 should increase the accuracy of secondary structure prediction.

Applying the PHD Scheme

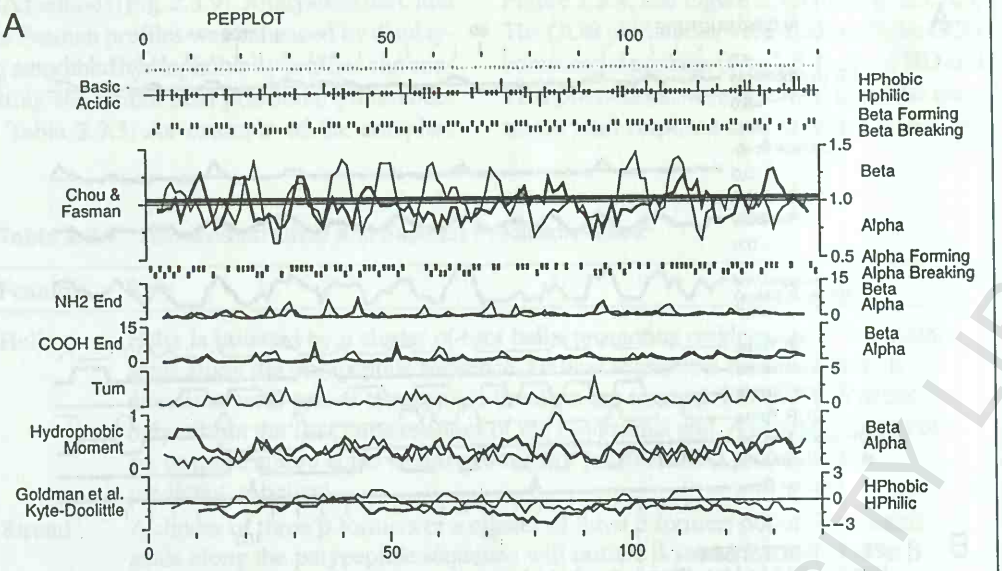
Access to the PHD scheme is available only through electronic mail or the WWW (World Wide Web). The Internet address for this service is given in Table 2.3.6. To perform sequence analysis with the PHD scheme, one merely needs to supply the sequence to the EMBL Predictprotein server. Additionally, one can submit prealigned sequences or a defined set of homolog sequences to be used in the predictive scheme. If a single sequence is submitted, the scheme will search the sequence database, and, if homologs are found, the sequence prediction will be performed on the multiple sequence alignments (Fig. 2.3.6). The output from the scheme consists of several pages describing the network and analysis, including the list of any homologous proteins along with the sequence alignment used in the analysis, the definition of accessibility, and a reliability index for the accuracy of prediction (Fig. 2.3.7). This method provides prediction of solvent accessibility as well as secondary structure. A comparison of this analysis to others described in this unit is described in the section on Comparison of Prediction Schemes.

Applying the PSA Scheme

The PSA scheme, like the PHD scheme, has been established on the Internet (Table 2.3.6). The electronic mail server, called Protein Sequence Analysis (PSA) System, accepts as input the amino acid sequence of a protein. After sequence analysis the server returns four

RGU IAS LIBRARY

A



B

Pos	Res	Alpha Stat	Alpha Ave	Beta Stat	Beta Ave	Alpha NH2	Alpha COOH	Beta NH2	Beta COOH	Turn	HPhob
1	A	1.42	0.94	0.83	0.96	0.43	0.17	0.29	0.56	0.30	-0.56
2	G	0.57	0.94	0.75	0.96	0.73	0.04	2.33	0.42	0.15	-0.33
3	W	1.08	0.97	1.37	1.14	0.57	0.25	0.47	0.76	0.28	-0.08
4	N	0.67	0.97	0.89	1.20	1.08	0.78	0.75	1.41	0.78	0.22
5	A	1.42	1.05	0.83	1.11	0.51	1.25	0.16	1.75	0.04	0.22
6	Y	0.69	0.86	1.47	1.13	0.25	0.31	1.24	1.27	0.45	-0.02
7	I	1.08	0.99	1.60	1.08	0.58	0.08	2.44	0.52	0.63	0.04
8	D	1.01	1.09	0.54	0.95	2.19	0.12	0.56	1.08	0.24	-0.18
9	N	0.67	1.19	0.89	1.02	1.98	0.95	0.09	2.40	0.03	-0.18
10	L	1.21	1.27	1.30	0.93	0.82	3.18	0.04	2.33	0.14	-0.18
11	M	1.45	1.11	1.05	0.79	0.31	1.52	0.31	0.63	1.41	-0.40
12	A	1.42	0.96	0.83	0.83	0.84	0.17	2.28	0.22	0.99	-0.41
13	D	1.01	0.78	0.54	0.92	2.45	0.13	1.21	0.25	1.04	-0.11
14	G	0.57	0.80	0.75	1.06	3.18	0.13	1.15	0.46	1.26	0.74
15	T	0.83	0.91	1.19	1.01	1.79	0.29	0.45	0.39	0.14	0.74
16	C	0.70	1.06	1.19	0.92	0.53	0.19	2.91	0.18	1.52	0.68
17	Q	1.11	1.24	1.10	0.83	0.27	0.13	2.86	0.10	0.17	-0.03
18	D	1.01	1.23	0.54	0.95	0.15	0.40	1.29	0.25	0.22	-0.03
19	A	1.42	1.25	0.83	1.24	0.12	0.71	0.30	0.76	0.03	0.27
20	A	1.42	1.03	0.83	1.22	0.14	3.45	0.11	2.52	0.09	-0.11
21	I	1.08	0.85	1.60	1.38	0.75	0.78	0.15	1.95	0.49	-0.40
22	V	1.06	0.87	1.70	1.17	1.03	0.56	0.18	4.19	0.57	-0.43
23	G	0.57	0.86	0.75	0.88	4.24	0.87	0.31	3.57	0.39	-1.00
24	Y	0.69	0.91	1.47	0.88	1.08	1.11	0.49	4.63	1.79	-0.76
25	K	1.16	0.88	0.74	0.65	1.39	0.50	1.50	0.74	0.51	-0.41

Figure 2.3.3 Secondary structure prediction for human profilin I using the method of Chou and Fasman as implemented by GCG:PepPlot algorithm. **(A)** Predicted secondary structure and **(B)** numerical output for residues 1 through 25.

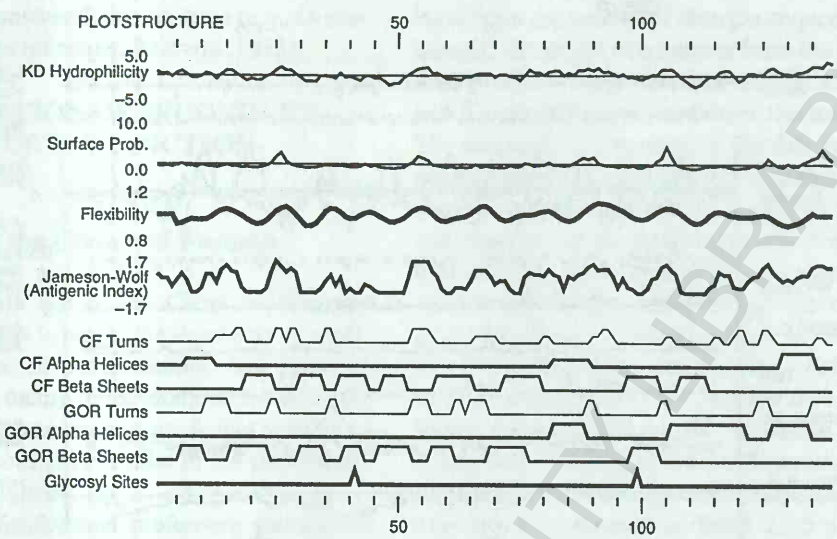
output files to the requester: a text file summarizing the tertiary class probabilities and most probable superclass; a file containing plots of the class and superclass probabilities; a file containing the secondary structure probability distributions displayed on a contour plot; and a file containing the secondary structure probability distributions on *x-y* plots. The secondary structure and protein class are easily determined from the combination of plots (Fig. 2.3.8) and can be compared with those suggested by other structural prediction methods.

COMPARISON OF PREDICTION SCHEMES

The secondary structures of hemoglobin (an all-α protein), interleukin 1β (an all-β protein), and profilin (an α/β protein) were predicted

DEAKIN UNIVERSITY LIBRARY

A



B

PEPTIDESTRUCTURE
Hydrophilicity (Kyte-Doolittle) averaged over a window of: 7
Surface Probability according to Emini
Chain Flexibility according to Karplus-Schulz
Secondary Structure according to Chou-Fasman
Secondary Structure according to Garnier-Osguthorpe-Robson
Antigenicity Index according to Jameson-Wolf

Pos	AA	GlycoS	HyPhil	SurfPr	FlexPr	CF-Pred	GORPred	AI-Ind ..
1	A	.	0.750	0.633	1.000	.	.	0.750
2	G	.	0.240	0.500	1.000	.	.	0.450
3	W	.	0.417	0.613	1.000	.	.	0.450
4	N	.	-0.286	0.426	1.000	.	.	-0.150
5	A	.	0.471	0.718	0.921	.	.	0.300
6	Y	.	0.914	1.099	0.927	.	B	0.750
7	I	.	0.243	0.563	0.946	h	B	0.300
8	D	.	-0.529	0.552	0.962	h	B	-0.600
9	N	.	-0.529	0.356	0.959	h	B	-0.600
10	L	.	-0.214	0.848	0.958	h	B	-0.300
11	M	.	0.486	0.502	0.960	h	T	0.700
12	A	.	0.086	0.451	0.983	h	T	0.700
13	D	.	-0.771	0.293	1.016	t	T	0.150
14	G	.	0.271	0.513	1.039	t	T	1.050
15	T	.	1.043	0.848	1.049	t	T	1.350
16	C	.	1.043	0.513	1.036	t	B	0.950
17	Q	.	0.286	0.523	1.018	T	B	0.850
18	D	.	-0.414	0.254	0.994	T	B	-0.200
19	A	.	-1.114	0.352	0.963	B	B	-0.600
20	A	.	-0.700	0.201	0.938	B	B	-0.600
21	I	.	-1.014	0.189	0.921	B	B	-0.600
22	V	.	-0.957	0.374	0.927	B	B	-0.600
23	G	.	-0.200	0.618	0.949	B	B	-0.300
24	Y	.	0.171	1.181	0.989	B	T	0.850
25	K	.	1.043	2.460	1.030	T	T	1.700
26	D	.	1.757	3.331	1.056	T	.	1.300
27	S	.	1.100	1.578	1.068	.	.	0.900
28	P	.	1.043	0.830	1.042	T	.	1.150
29	S	.	0.229	0.502	1.000	T	.	0.850
30	V	.	-0.529	0.378	0.951	B	B	-0.600
31	W	.	-1.243	0.182	0.908	B	B	-0.600
32	A	.	-1.243	0.210	0.906	B	B	-0.600
33	A	.	-1.300	0.279	0.929	B	B	-0.600
34	V	.	-0.143	0.531	0.975	B	B	-0.300
35	P	.	-0.171	0.759	1.029	T	.	0.250

Figure 2.3.4 Secondary structure prediction for human profilin I using the methods of Chou and Fasman and GOR as implemented by GCG: PeptideStructure algorithm and plotted using PlotStructure utility. **(A)** Predicted secondary structure and **(B)** numerical output for residues 1 through 35.

BCU IASI/CENTRAL UNIVERSITY LIBRARY

with the Chou and Fasman, GOR, PHD, and PSA methods (Fig. 2.3.9). Analysis of the Chou and Fasman profiles was enhanced by displaying smoothed hydrophobicity profiles and consulting the amino acid positional preferences of Table 2.3.3. An example of the complete analysis of profilin is presented in Figure 2.3.3, Figure 2.3.4, and Figure 2.3.5 (or Fig. 2.3.2C). The GOR predictions were made with the GCG commercial package (Fig. 2.3.4). The PHD and PSA predictions were obtained from the electronic mail response to a structure query; ex-

Table 2.3.4 Abbreviated Chou and Fasman Prediction Rules[a]

Feature	Rule
Helix	Helix is initiated by a cluster of four helix-promoting residues out of six amino acids along the polypeptide sequence. Helical segment is extended in both directions until sets of tetrapeptide breakers are reached. Prolines can occur only within the first three residues of the N-terminal end. A segment length of six or more amino acids with α probability greater than β probability is predicted as helical.
Strand	A cluster of three β formers or a cluster of three β formers out of five amino acids along the polypeptide sequence will initiate β-strand formation. The β strand is extended in both directions until terminated by a set of tetrapeptide breakers. A segment of three or more amino acids with β probability greater than α probability is predicted as β strand.
Turns	Turn probability is calculated as peaks of probability that are greater than α or β probability. The β turns are also identified by $Pt > 0.75 \times 10^{-4}$, where $Pt = f_i \times f_{i+1} \times f_{i+2} \times f_{i+3}$.
Overlap	For a region containing overlapping α helix and β strand, the prediction would be for helix if the α probabilities are greater than β probabilities, and strand if β probabilities exceed α probabilities.

[a]For the complete set of prediction rules, see Chou and Fasman (1974a,b, 1978a,b).

Table 2.3.5 Secondary Structure Correlations

Structure	Characteristics
β strands	Often correspond to peaks of hydrophobicity β-branched amino acids (Val, Ile, Thr) are typical Positively charged residues are relatively abundant Negatively charged residues (Glu) virtually never occur Average residue length is 7, resulting in short and acute peaks in hydrophobicity plots (see *UNIT 5.2*) β bulges may occur in the edge strands. Their signature is small amino acids and glycines (e.g., Gly-Xxx-Gly sequence in immunoglobulin V fold, strand G).
α helices	Hydrophobicity peaks are longer, with a fine structure (teeth). Ideally, fine structure shows $n + 4$ periodicity. Glu, Ala, Leu, are typical Amino acid preferences exist at the N-cap and C-cap positions (see Table 2.3.3) Prolines can occur in the first turn of helices in globular proteins, and most frequently occupy the N-cap + 1 position Prolines, which "kink" the helices, are overrepresented in transmembrane helices Glycines frequently occur at the C-cap position Hydrophobic residues often occupy positions N-cap + 4 and C-cap − 4 Helices may be completely buried and not amphipathic

BCUJAB CENTRAL UNIVERSITY LIBRARY

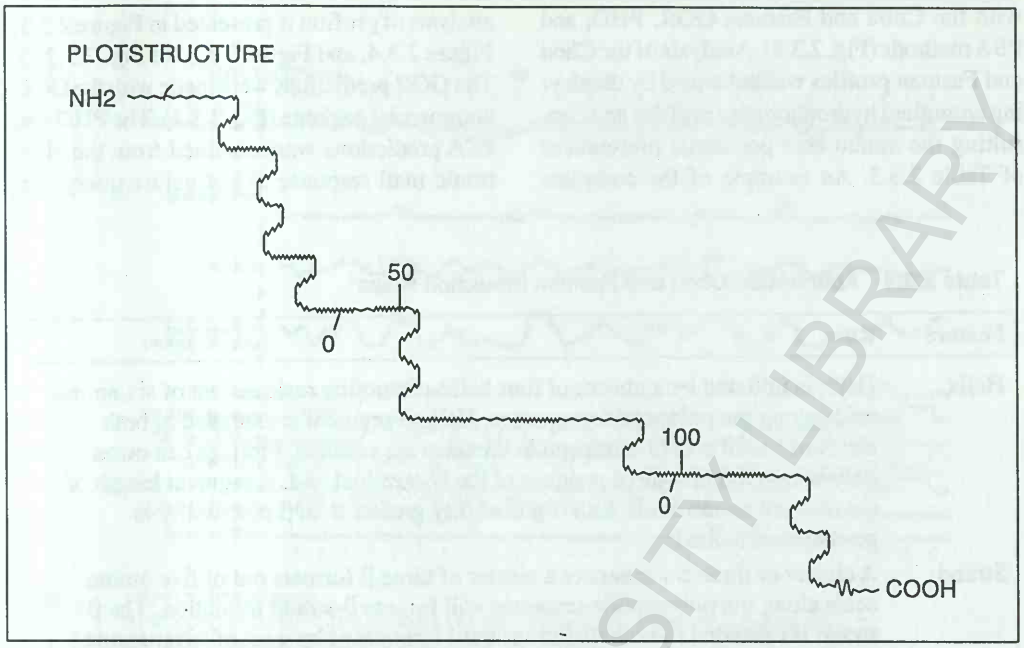


Figure 2.3.5 Sketch of the human profilin secondary structure as predicted in Figure 2.3.3 and Figure 2.3.4 by Chou-Fasman method.

amples of the output for the profilin secondary structure prediction are provided in Figure 2.3.6, Figure 2.3.7, and Figure 2.3.8.

Figure 2.3.9A summarizes the results of secondary structure prediction for the β subunit of human hemoglobin. Immediately below the amino acid sequence is the location of the eight helices found in the three-dimensional structure of hemoglobin (Brookhaven Protein Data Bank entry 2hhb; Fermi, 1975). The Chou and Fasman scheme predicts most of the helical residues in helices 1, 2, 3, 4, 6, 7, and 8. Helices 6 and 7 are predicted to be longer than they actually are, and helices 2 and 3 are not distinguished as independent helices. A turn is predicted in place of the fifth helix. The GOR method predicts most of the residues for helices 1, 2, 5, 6, and 8, but it predicts a turn for helix 3 as well as a turn and a strand for helix 7. The PHD program found over 400 sequence homologs of hemoglobin-β, with one of the sequences being the submitted sequence. The PHD prediction identified a homolog with a known three-dimensional structure in the protein database and also suggested structure prediction by homology modeling. The prediction by PHD did quite well, accounting for seven of the eight helices and missing only the small helix 4. The PSA procedure suggested the α/β class as the most probable. Most of the helical residues in helices 2, 3, 5, 6, and 8 and part of helix 7 were correctly predicted. Helix 1 was missed by the PSA method, however, and the C- and N-terminal residues of helices 2 and 3, respectively, were predicted to be in sheet structure.

Figure 2.3.9B summarizes the results of secondary structure prediction for interleukin 1β. IL-1β contains 12 β strands connected by intervening turns (Brookhaven Protein Data Bank entry 1i1b; Finzel et al., 1989). The Chou and Fasman method predicted most of the residues in 8 of the 12 strands in the β sheet (strands 1, 2, 4, 5, 6, 8, 11, and 12). Helical regions were predicted for β strands 3 (with the adjacent turn) and 7. The Chou and Fasman method correctly predicted several of the turns. The GOR method correctly predicted 4 of the 12 β strands, but it predicted helical regions for β strands 3 (with the adjacent turn), 5, 6, and 8. Interestingly, both the Chou and Fasman and GOR methods predict helix for the third strand. This may be due to the negatively charged glutamate, which is generally considered a strand breaker (see Table 2.3.1). For IL-1β, the PHD output listed the 10 sequence homologs that were used in the structure prediction and also identified a known three-dimensional structure in the protein database. The PHD method correctly predicted most of the residues in the 12 β strands; however, a deficiency in the method is that it does not predict the location of possible turns. The PSA method did not successfully predict either the class or the secondary structure of IL-1β: the method pre-

RCU IASRI LIBRARY UNIVERSITY LIBRARY

Table 2.3.6 Commercial Packages for Sequence Analysis

Software	Platform[a]	Vendor or Internet location[b]
Sybyl/Biopolymer[c]	SGI[d]	Tripos Associates
InsightII/Homology[c]	SGI[d]	Biosym Technologies
GCG	VAX	Genetics Computer Group
MacVector	Macintosh	International Biotechnologies
Geneworks	Macintosh	IntelliGenetics/Betagen
PC/Gene	IBM-compatible	IntelliGenetics/Betagen
DNASTAR	IBM-compatible or Macintosh	DNASTAR
Internet locations for sequence analysis		
PHD method[e]		EMBL Predictprotein Server predictprotein@embl-heidelberg.de predict-help@embl-heidelberg.de http://www.embl-heidelberg.de/predictprotein/predict protein.html
PSA method[e]		Protein Sequence Analysis Server psa-request@darwin.bu.edu

[a]Software may run on other platforms in addition to these listed here. For a comprehensive list, contact the appropriate vendor.

[b]For suppliers addresses, see *SUPPLIERS APPENDIX*.

[c]Biopolymer and Homology are optional modules that can be purchased to run with Sybyl and InsightII, respectively.

[d]Silicon Graphics Inc.

[e]PHD and PSA methods are currently accessible only through the Internet; thus, results can be accessed by any platform.

dicted the protein fold class to be all-α rather than all-β, and therefore incorrectly predicted α structure rather than β structure. It is noteworthy that several of the predicted α helices correspond to the location of β strands, namely, those of β strands 2, 3, 4, 5, 7, 9, 11, and 12. Further testing is required to determine whether there are any inherent problems with the predictions of all-β proteins by the PSA method.

Figure 2.3.9C summarizes the results of secondary structure prediction for human profilin, an α/β protein whose high-resolution structure (Metzler et al., 1993, 1995) was not publicly known at the time of testing. Profilin is composed of a central seven-stranded, antiparallel β sheet and four helices (Fig. 2.3.10). Seven turns that connect the helical and sheet structural elements in the structure have been identified. Figure 2.3.2C, Figure 2.3.3A, and 2.3.4A show typical Chou and Fasman profiles as determined with the programs of Novotny (Novotny and Auffray, 1984), PepPlot (GCG, 1994), and Plotstructure (GCG, 1994). Figure 2.3.3B shows the numerical output from the modified Chou and Fasman prediction programs (PepPlot), and Figure 2.3.4B shows the automated prediction by the Peptidestructure program using the Chou and Fasman and the GOR methods. A schematic representation of the predicted protein secondary structure is shown in Figure 2.3.5.

The Chou and Fasman method predicted six of the seven β strands and the two terminal helices of profilin (Fig. 2.3.8C). This method also predicted correct locations of several of the turns located between the strands. In comparison, the GOR method performed significantly worse, predicting only four of seven β strands and neither of the terminal helices. The PHD method identified six sequence homologs of profilin (Fig. 2.3.6) that it used in the secondary structure prediction (Fig. 2.3.6). The method correctly predicted all seven β strands and the two terminal helices. It did, however, predict additional β strands for the interior helices but did not predict the locations of any turns (owing to a deficiency in the method, as stated previously). The PSA method predicted the profilin sequence to be a member of the all-β protein fold class (Fig. 2.3.8). The method

BARNARD UNIVERSITY LIBRARY

```
## PROTEINS : EMBL/SWISSPROT identifier and alignment statistics
   NR.    ID         STRID   %IDE %WSIM IFIR ILAS JFIR JLAS LALI NGAP LGAP LSEQ2 ACCNUM
PROTEIN
    1 : pro1_human         1.00  1.00    1   139    1   139   139    0    0   139   P07737
PROFILINI.
    2 : prof_mouse         0.96  0.97    1   139    1   139   139    0    0   139   P10924
PROFILIN.
    3 : prof_bovin         0.95  0.97    1   139    1   139   139    0    0   139   P02584
PROFILIN.
    4 : pro2_human         0.62  0.71    1   139    1   139   139    0    0   139   P35080
PROFILINII.
    5 : prof_varv          0.32  0.48    1   137    2   131   130    2    7   133   P33828
PUTATIVEPROFILIN.
    6 : prof_vaccv         0.31  0.47    1   137    2   131   130    2    7   133   P20844
PUTATIVEPROFILIN.

## ALIGNMENTS    1 -    6
   SeqNo  PDBNo AA STRUCTURE BP1 BP2 ACC  NOCC   VAR   ....:....1....:....2....:....3....:...
.4....:....5....:....6....:....7
       1        A  U              0   0   0     6     0   AAAAAA
       2        G  U              0   0   0     6    30   GGGGEE
       3        W  U              0   0   0     6     0   WWWWWW
       4        N  U              0   0   0     6    35   NNNQHH
       5        A  U              0   0   0     6    50   AAASKK
       6        Y  U              0   0   0     6    41   YYYYII
       7        I  U              0   0   0     6     5   IIIVII
       8        D  U              0   0   0     6    15   DDDDEE
       9        N  U              0   0   0     6    29   NSNNDD
      10        L  U              0   0   0     6    21   LLLLII
      11        M  U              0   0   0     6    53   MMMMSS
      12        A  U              0   0   0     6    56   AAACKK
      13        D  U              0   0   0     6    24   DDDDNN
      14        G  U              0   0   0     6    33   GGGGNN
      15        T  U              0   0   0     6    51   TTTCNK
      16        C  U              0   0   0     6    47   CCCCFF
      17        Q  U              0   0   0     6    24   QQQQEE
      18        D  U              0   0   0     6     6   DDDEDD
      19        A  U              0   0   0     6     0   AAAAAA
      20        A  U              0   0   0     6     0   AAAAAA
      21        I  U              0   0   0     6     0   IIIIII
      22        V  U              0   0   0     6     0   VVVVVV
      23        G  U              0   0   0     6    24   GGGGDD
      24        Y  U              0   0   0     6     0   YYYYYY
      25        K  U              0   0   0     6    27   KKKCKK
```

Figure 2.3.6 Sequence homologs and a portion of the sequence alignment statistics used in secondary structure prediction for human profilin I with PHD, the EMBL neural network method. Secondary structure prediction given in Figure 2.3.7. For description of column labels see Rost and Sander (1993a,b) and EMBL Predictprotein Server (address given in Table 2.3.6).

did correctly predict the location of five of the β strands, but it predicted additional β strands for the amino and carboxyl termini and the second interior helices. The PSA method correctly predicted several of the turns located between the β strands, as well.

In summary, both the PHD and the interactive Chou and Fasman method performed well, correctly predicting most secondary structural elements (see Table 2.3.7). The interactive Chou and Fasman method generally outperformed the GOR method. However, it must be pointed out that in predicting the secondary structure with the Chou and Fasman method, the correlations contained in Table 2.3.3 and Table 2.3.5 and the Chou and Fasman rules contained in Table 2.3.4 facilitated both locating helix ends and discriminating between helix and strand when ambiguity existed. In contrast, a strict interpretation of the GOR results was used. It is also possible to interpret the GOR results interactively; this may equalize the predictive power of the two methods.

In testing the PHD method, we were unable

RCU IAS / CENTRAL UNIVERSITY LIBRARY

Table 2.3.7 Summary of Secondary Structure Prediction

Structure	Method (%)[a]			
	C-F	GOR	PHD	PSA
Hemoglobin	78	62	86	60
Interleukin 1β	53	41	67	—
Profilin	60	44	87	48

[a]Number of amino acids correctly predicted, divided by the total number of amino acids known to be in helices, sheets, or turns. No value is provided for the PSA analysis of interleukin 1β because the class predicted was all-α.

```
protein:        2a20            length      139

                  ....,....1....,....2....,....3....,....4....,....5....,....6
          AA      |AGWNAYIDNLMADGTCQDAAIVGYKDSPSVWAAVPGKTFVNITPAEVGVLVGKDRSSFYV|
          PHD sec |  HHHHHHHHH        EEEEE    EEEEEE   EEEEE   EEEEEEEE   EEEE|
          Rel sec |9656799887337973560899642586359862574486654817779997268747 87|
 detail:
          prH sec |0267788887631002100000000001100000000000000001000000000000000
          prE sec |0000000100000012114898764211568875212687773047889897521167 88
          prL sec |9721100011368975674100135686220123686312226841110001478831 11
 subset:  SUB sec |LLHHHHHHHH..LLL.LL.EEEE..LLL.EEEE.LL..EEEE.L.EEEEEEE.LLL.EEE|

 ACCESSIBILITY
 3st:     P_3 acc |eebeebbebbbeeeebebbbbbbbee eebbbbbbeebbbeb be bbbbbeee eebbb|
 10st:    PHD acc |9806700600077770600000007757700000077000705074000007775760 00
          Rel acc |8422318104045250128977404406475973045151462130842473551413 24
 subset:  SUB acc |ee....b..b.ee.e...bbbbb.ee.eebbbb..ee.b.eb....bb.bb.ee.e...b|

                  ....,....7....,....8....,....9....,....10...,....11...,....1
          AA      |NGLTLGGQKCSVIRDSLLQDGEFSMDLRTKSTGGAPTFNVTVTKTDKTLVLLMGKEGVHG|
          PHD sec |EEEEE   EEEEEEEEE     EEEEEE          EEEEEE   EEEEEE       |
          Rel sec |3246277179999752245893159976257999883139999755258888826787 67|
 detail:
          prH sec |0000000000000001110000110000010000000000000011211000000011100
          prE sec |6567511489999865422003468877411000113458999821268888841000 11
          prL sec |3422478510000123367885420011568989886530000167410000057888 77
 subset:  SUB sec |...E.LL.EEEEEEE...LLL..EEEEE.LLLLLLL...EEEEELL.EEEEEE.LLLLLL|

 ACCESSIBILITY
 3st:     P_3 acc |bbbebbeeebbbb bbbbeebebbbeb beeeeebbbbebbbbebbebbbbbbbeeebee|
 10st:    PHD acc |0006007760000500007707000605077797000060000600700000007760 77
          Rel acc |2011004428083116404405427251374545110117381110559976844500 44
 subset:  SUB acc |......ee.b.b...bb.ee.eb.b.b..eeeee.....b.b....ebbbbbbbee..ee|

                   2...,....13...,....14...,....15...,....16...,....17...,....1
          AA      |GLINKKCYEMASHLRRSQY|
          PHD sec |   HHHHHHHHHHHHHHH   |
          Rel sec |8424899999999972589
 detail:
          prH sec |0146889999999985200
          prE sec |0210000000000000000
          prL sec |8632100000000014689
 subset:  SUB sec |L...HHHHHHHHHHH.LLL|

 ACCESSIBILITY
 3st:     P_3 acc |eebeeeb ebbeeb eeee|
 10st:    PHD acc |7606770570076057789
          Rel acc |4131359142931413469
 subset:  SUB acc |e....eb.e.b..b..eee|
```

Figure 2.3.7 Secondary structure prediction for human profilin I with PHD, the EMBL neural network method. Sequence homologs and a portion of the sequence alignment used in the analysis are given in Figure 2.3.6.

DUKE UNIVERSITY MEDICAL CENTER LIBRARY

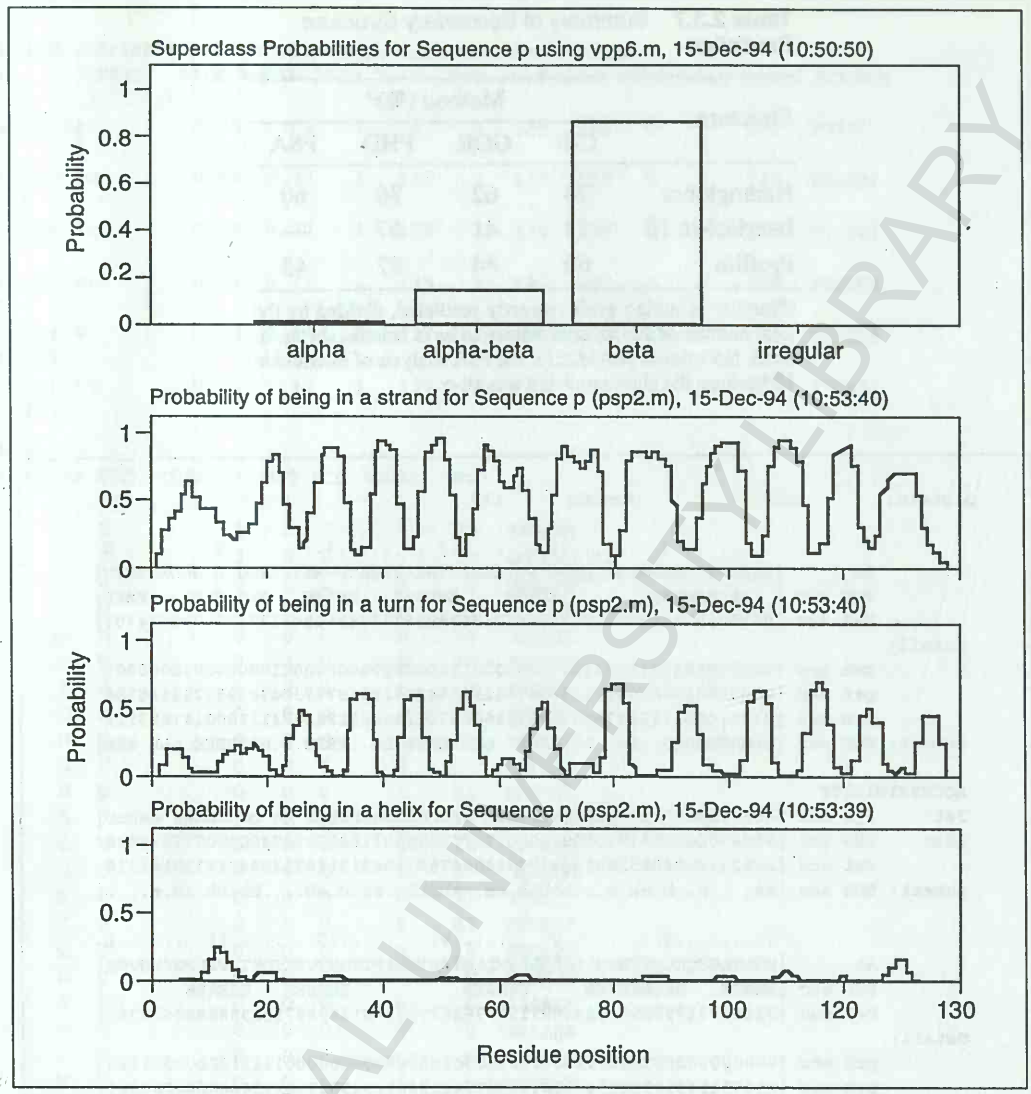


Figure 2.3.8 Class and secondary structure predictions for human profilin I with the PSA method.

to identify any protein whose structure is included in the Brookhaven Protein Data Bank for which there are no sequence homologs in sequence databases. If such a sequence and structure were available, the comparison of the predictive methods might have been less biased. However, this point highlights an advantage of the PHD method. Sequence homologs with known three-dimensional structures often help achieve an accurate prediction of secondary structure. The availability of homologous amino acid sequences imparts a significant predictive advantage to the PHD method in many cases.

RECOMMENDATIONS AND THE FUTURE

On the basis of the examples presented here and our experience in general, a combination of the Chou and Fasman interactive sequence analysis (which requires interpretation of the output data) and PHD neural network predictions (which are automatic) is recommended as the most accurate and preferable procedure.

As three-dimensional protein structural databanks grow, the future of protein structure prediction may lie with methods that "thread" the polypeptide sequence onto known motifs and evaluate the resulting model by pseudoenergetic analysis (e.g., PROSA; Sippl and Weitckus, 1992). Such programs currently exist, and successful fold identifications are being reported. However, the probability of a successful sequence-structure match is currently 10% (Sippl and Weitckus, 1992).

CONCLUSIONS

1. Empirically derived rules can aid sequence prediction using the GOR and Chou and Fasman statistical methods. Summaries of

RGU IAIN B. BANKS CENTRAL UNIVERSITY LIBRARY

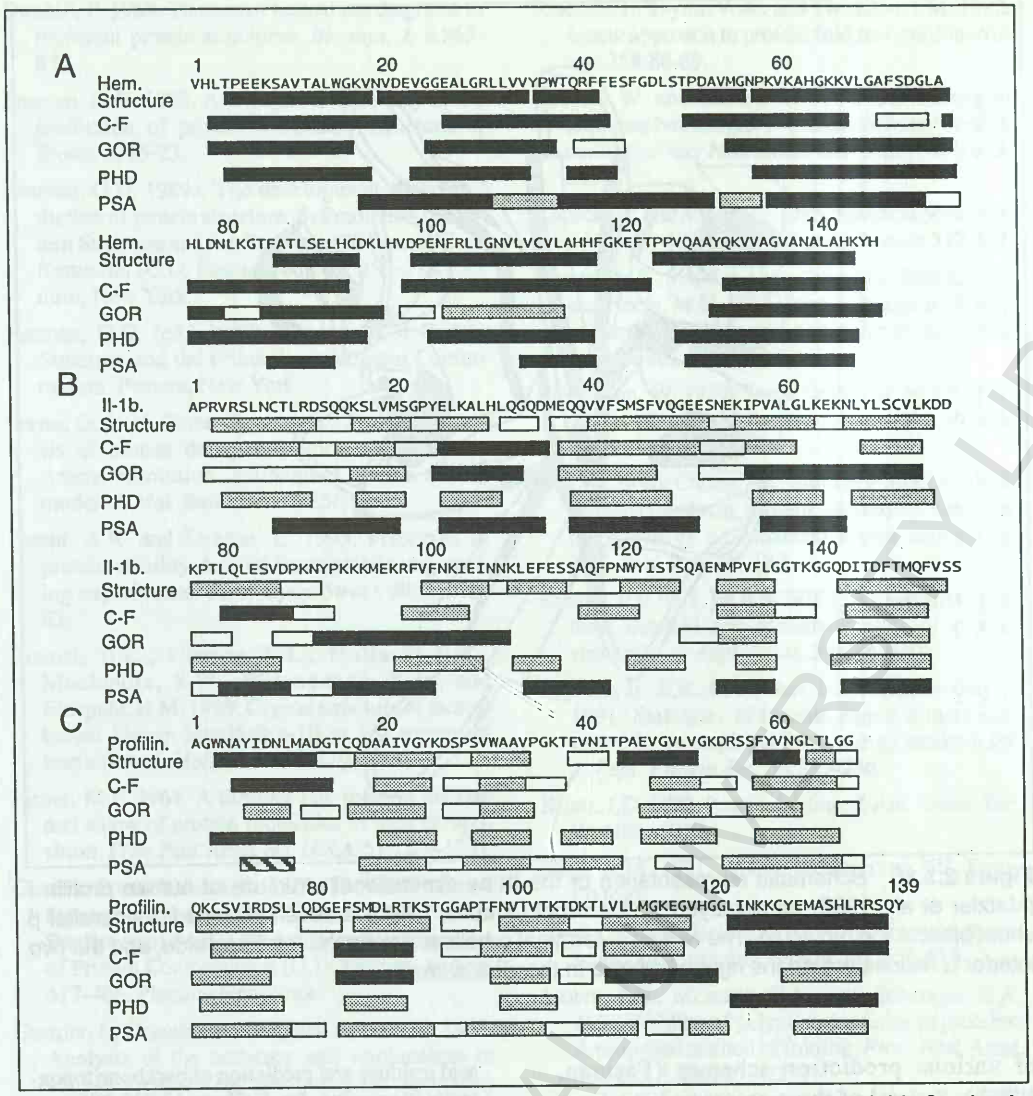


Figure 2.3.9 Summary of the secondary structure predicted for (**A**) human hemoglobin β-subunit, (**B**) human interleukin 1β, and (**C**) human profilin I. Each panel lists, from top to bottom, the amino acid sequence, the secondary structure determined experimentally, and the secondary structure predicted by each of the following: C-F, the Chou and Fasman method; GOR, the GOR method; PHD, the EMBL neural network method; and PSA, the state-space method. The location of helices (solid bars), turns (open bars), and sheet structure (hatched bars) are indicated.

these correlations (rules) are contained in Tables 2.3.4 and 2.3.5.

2. Methods that are open to interpretation do not guarantee the same result when applied by different users.

3. The PHD and PSA methods allow no input from the user, other than the sequence of interest, and thus are the easiest to use. For the other methods described here, the user can interactively participate in the prediction process.

4. If sequence homologs and, in particular, three-dimensional data are available, the PHD method gives reliable secondary structure predictions.

5. The PSA method was found to be less accurate than the other predictive schemes tested.

ACCESSIBILITY OF SOFTWARE

Several prediction schemes have been incorporated into commercial packages (e.g., GCG) that also contain algorithms for additional sequence analysis. Other packages have interactive graphics modules that combine sequence analysis with molecular displays (Tripos and Biosym); however, these packages do not easily incorporate additional user input. There have been numerous reports in the literature describing computerized implementation

BOJANGLES MEMORIAL LIBRARY

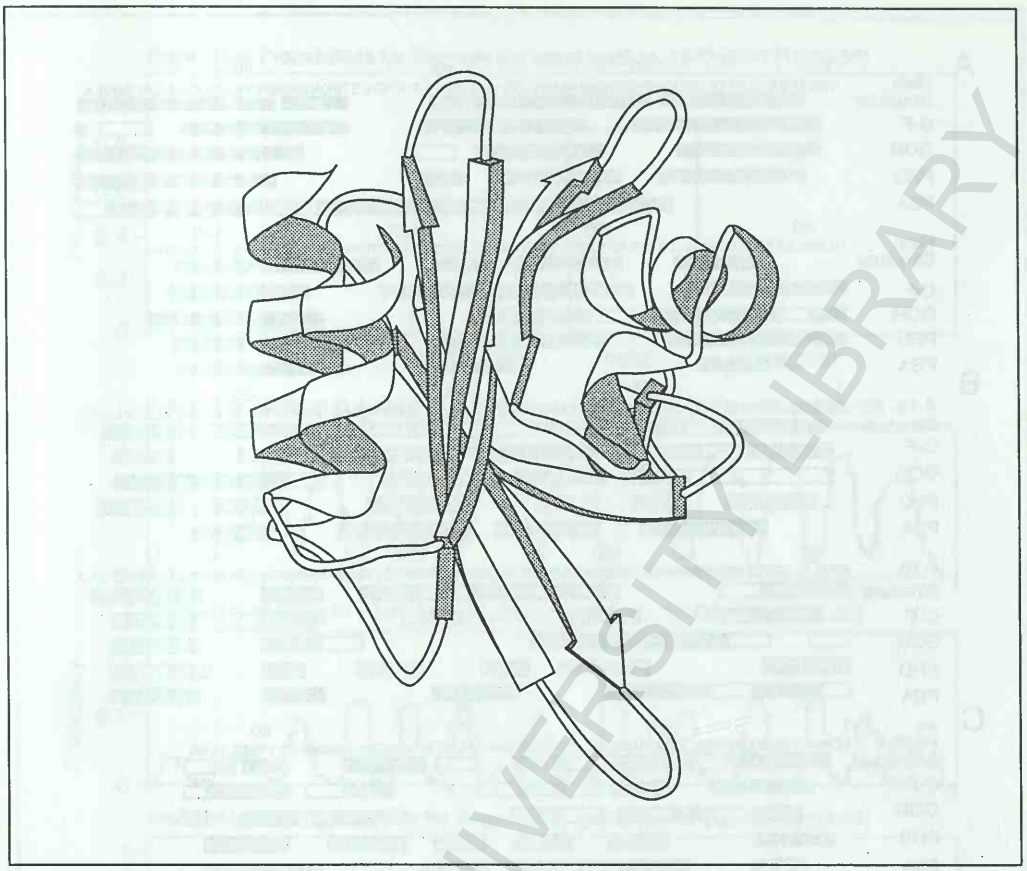


Figure 2.3.10 Schematic representation of the three-dimensional structure of human profilin I (Metzler et al., 1995) as displayed by Molscript (Kraulis, 1991). A seven-stranded antiparallel β sheet bisects the molecule. The N- and C-terminal α helices are on the left-hand side, and the two interior α helices are on the right-hand side in the view shown.

of various prediction schemes (Fasman, 1989b). Several of these programs have been published in a variety of programming languages. There are also several sites available on the Internet for sequence analysis. A listing of some of the commercial and Internet programs is contained in Table 2.3.6.

Literature Cited

Bajorath, J., Stenkamp, R., and Aruffo, A. 1993. Knowledge-based model building of proteins: Concepts and examples. *Protein Sci.* 2:1798-1810.

Baldwin, J.M. 1993. The probable arrangement of helixes in G protein–coupled receptors. *EMBO J.* 12:1693-1703.

Blundell, T.L. and Johnson, L.N. (eds.) 1976. Protein Crystallography. Academic Press, London.

Branden, C. and Tooze, J. 1991. Introduction to Protein Structure. Garland Publishing, New York.

Brookhaven. 1994. Protein Data Bank. Brookhaven National Laboratory, Upton, N.Y.

Burgess, A.W., Ponnuswamy, P.K., and Scheraga, H.A. 1971. Analysis of conformations of amino acid residues and prediction of backbone topography in proteins. *Isr. J. Chem.* 12:239-286.

Chou, P.Y. 1989. Prediction of protein structural class from amino acid compositions. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 549-586. Plenum, New York.

Chou, P.Y. and Fasman, G.D. 1974a. Conformational parameters for amino acids in helical, beta sheet and random coil regions calculated from proteins. *Biochemistry* 13:211-222.

Chou, P.Y. and Fasman, G.D. 1974b. Prediction of protein conformation. *Biochemistry* 13:223-245.

Chou, P.Y. and Fasman, G.D. 1977. β-turns in proteins. *J. Mol. Biol.* 115:135-175.

Chou, P.Y. and Fasman, G.D. 1978a. Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47:251-276.

Chou, P.Y. and Fasman, G.D. 1978b. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 47:45-148.

DeGrado, W.F. 1988. Design of peptides and proteins. *Adv. Protein Chem.* 39:51-124.

Dill, K.A. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133-7155.

BCU IASL - ETHAL UNIVERSITY LIBRARY

Dunhill, P. 1968. The use of helical net diagrams to represent protein structures. *Biophys. J.* 8:865-875.

Fasman, G.D. 1985. A critique of the utility of the prediction of protein secondary structure. *J. Biosci.* 8:15-23.

Fasman, G.D. 1989a. The development of the prediction of protein structure. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 193-316. Plenum, New York.

Fasman, G.D. (ed.) 1989b. Prediction of Protein Structure and the Principles of Protein Conformation. Plenum, New York.

Fermi, G. 1975. Three-dimensional Fourier synthesis of human deoxyhaemoglobin at 2.5 angstroms resolution, refinement of the atomic model. *J. Mol. Biol.* 97:237-256.

Fersht, A.R. and Serrone, L. 1993. Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* 3:75-83.

Finzel, B.C., Clancy, L.L., Holland, D.R., Muchmore, S.W., Watenpaugh, K.D., and Einspahr, H.M. 1989. Crystal structure of recombinant human interleukin-1β at 2.0 angstroms resolution. *J. Mol. Biol.* 209:779.

Fisher, M.F. 1964. A limiting law relating the size and shape of protein molecules to their composition. *Proc. Natl. Acad. Sci. U.S.A.* 51:1285-1291.

Garnier, J. and Robson, B. 1989. The GOR method for predicting secondary structure in proteins. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 417-466. Plenum, New York.

Garnier, J., Osguthorpe, D.J., and Robson, B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97-120.

Genetics Computer Group. 1994. (GCG) Wisconsin Sequence Analysis Package Program Manual. Genetics Computer Group, Inc., Madison, Wis.

Goodman, E.M. and Kim, P.S. 1991. Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry* 30:11615-11620.

Greer, J. 1990. Comparative modeling: Application to the family of the mammalian serine proteases. *Proteins Struct. Funct. Genet.* 7:317-334.

Guzzo, A.V. 1965. The influence of amino acid sequence on protein structure. *Biophys. J.* 5:809-822.

Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262:1401-1407.

Harper, E.T. and Rose, G.D. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry* 32:7605-7609.

Jones, D.T., Taylor, W.R., and Thornton, J.M. 1992. A new approach to protein fold recognition. *Nature* 358:86-89.

Kabsch, W. and Sander, C. 1984. On the use of sequence homologies to predict different conformations. *Proc. Natl. Acad. Sci. U.S.A.* 81:1075-1078.

Kabsch, W. and Sander, C. 1985. Identical pentapeptides with different backbones. *Nature* 317:207.

Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M., and Hecht, M.H. 1993. Protein design by binary patterning of polar and non-polar amino acids. *Science* 262:1680-1685.

Kauzmann, W. 1959. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14:1-63.

Kneller, D.G., Cohen, F.E., and Largridge, R. 1990. Improvements in protein secondary structure prediction by an enhanced neural network. *J. Mol. Biol.* 214:171-182.

Kraulis, P.J. 1991. MOLSCRIPT: A program to produce detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24:946-950.

Krystek, Jr., S.R., Bruccoleri, R.E., and Novotny, J. 1991. Stabilities of leucine zipper dimers estimated by an empirical force energy method. *Int. J. Pept. Protein Res.* 38:229-236.

Kuntz, I.D. 1972. Protein folding. *J. Am. Chem. Soc.* 94:4009-4012.

Lesk, A.M. 1991. Protein architecture. IRL Press, Oxford.

Levitt, M. and Chothia, C. 1976. Structural patterns in globular proteins. *Nature* 261:552-558.

Lewis, P.N., Momany, F.A., and Scheraga, H.A. 1971. Folding of polypeptide chains in proteins: A proposed method of folding. *Proc. Natl. Acad. Sci. U.S.A.* 68:2293-2297.

Lim, V.I. 1974. Algorithms for prediction of α helices and β-structural regions in globular proteins. *J. Mol. Biol.* 88:873-894.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D.K., DeGrado, W.F., and Eisenberg, D. 1993. Crystal structure of a synthetic triple-stranded α-helical bundle. *Science* 259:1288-1293.

Metzler, W.J., Constantine, K., Friedrichs, M.S., Bell, A., Ernst, E., Lavoie, T.B., and Mueller, L. 1993. Characterization of the three-dimensional structure of human profilin: ^{1}H, ^{13}C and ^{15}N assignments and global folding pattern. *Biochemistry* 32:13818-13829.

Metzler, W.J., Farmer, B.T., II, and Mueller, L. 1995. Refined solution structure of human profilin I. *Protein Sci.* 4:450-459.

Nagano, K. 1973. Logical analysis of the mechanism of protein folding. I. Prediction of helices, loops and β-structures from primary structure. *J. Mol. Biol.* 75:401-420.

BOSTON SPECTRA DIGEST LIBRARY

Novotny, J. and Auffray, C. 1984. A program for prediction of protein secondary structure from nucleotide sequence data: Application to histocompatability antigens. *Nucl. Acids Res.* 12:243-255.

Oppenheimer, N.J. and James, T.L. (eds.) 1989. Nuclear magnetic resonance (Part B). *Methods Enzymol.* Volume 177.

O'Shea, E.K., Rutkowski, R., and Kim, P.S. 1992. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* 68:699-708.

Pauling, L. and Corey, R.B. 1951. The pleated sheet, a new layer of configuration of polypeptide chains. *Proc. Natl. Acad. Sci. U.S.A.* 37:251-256.

Pauling, L., Corey, R.B., and Branson, H.R. 1951. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* 37:205-211.

Presta, L.G. and Rose, G.D. 1988. Helix signal in proteins. *Science* 240:1632-1641.

Prevelige, P., Jr. and Fasman, G.D. 1989. Chou-Fasman prediction of the secondary structure of proteins. *In* Prediction of Protein Structure and the Principles of Protein Conformation (G.D. Fasman, ed.) pp. 391-416. Plenum, New York.

Prothero, J.W. 1966. Correlation between the distribution of amino acids in alpha helices. *Biophys. J.* 6:367-370.

Ptitsyn, O.B. and Finkelstein, A.V. 1970. *Biofizika* 15:757-768.

Quinn, N. and Sejnowski, T.J. 1988. Predicting the secondary structure of globular proteins using neural network models. *J. Mol. Biol.* 202:865-884.

Richardson, J.S. 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34:167-339.

Richardson, J.S. and Richardson, D.C. 1988. Amino acid preferences for specific locations at the ends of α-helices. *Science* 240:1648-1652.

Robson, B. and Pain, R.H. 1971. Analyses of the code relating sequence to conformation in proteins: Possible implication for the mechanism of formation of helical regions. *J. Mol. Biol.* 58:237-259.

Rost, B. and Sander, C. 1993a. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc. Natl. Acad. Sci. U.S.A.* 90:7558-7562.

Rost, B. and Sander, C. 1993b. Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.* 232:584-599.

Roy, S. and Rose, G.D. 1980. Hydrophobic basis of packing in globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* 8:4643-4647.

Sander, C. and Schneider, R. 1991. Database of homology-derived structures and the structural meaning of sequence alignment. *Proteins Struct. Funct. Genet.* 9:56-68.

Schiffer, M. and Edmundson, A.B. 1967. The use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys. J.* 7:121.

Schulz, G.E. 1988. A critical evaluation of methods for prediction of protein secondary structures. *Annu. Rev. Biophys. Chem.* 17:1-21.

Schulz, G.E., Barry, C.D., Friedman, J., Chou, P.Y., Fasman, G.D., Finkelstein, A.V., Lim, V.I., Ptitsyn, O.B., Kabat, E.A., Wu, T.T., Levitt, M., Robson, B., and Nagano, K. 1974. Comparison of predicted and experimentally determined secondary structure of adenyl kinase. *Nature* 250:140-142.

Schulz, G.E. and Schirmer, R.H. 1979. Principles of Protein Structure. Springer-Verlag, New York and Heidelberg.

Sippl, M.J. and Weitckus, S. 1992. Detection of native-like models for amino acid sequences of unknown three-dimensional structure in a data base of known protein conformations. *Proteins Struct. Funct. Genet.* 13:258-271.

Stultz, C.M., White, J.V., and Smith, T.F. 1993. Structural analysis based on state-space modeling. *Protein Sci.* 2:305-314.

Wuthrich, K. 1986. NMR of Proteins and Nucleic Acids. John Wiley & Sons, New York.

Wyckoff, H.W., Hirs, C.H.W., and Timasheff, S.N. (eds.) 1985. Diffraction methods for biological molecules (Parts A and B). *Methods Enzymol.* Volumes 114 and 115.

Key Reference

Fasman, 1989b. See above.

Provides extensive descriptions and a historical perspective describing methods for protein structure prediction.

Contributed by Stanley R. Krystek, Jr., William J. Metzler, and Jiri Novotny
Bristol-Myers Squibb Pharmaceutical Research Institute
Princeton, New Jersey

Internet Basics

The Internet, more popularly known as the "information superhighway," has become a very pervasive influence in everyday life. Its power is most strongly seen among scientists, as more and more information is made available through "the Net," whether it be sequences, experimental data, chromosome maps, or whole journal articles. The aim of this unit is to provide a brief overview of what the Internet is, what kinds of information can be found on the Net, and how to use Internet tools to find and gather that information. The focus will be on practical examples, beginning with electronic mail and following the Internet timeline through to the development and use of the World Wide Web.

Despite the impression that the Internet is a single entity, it is actually a network of networks, composed of over 20,000 interconnected local and regional networks in more than 100 countries. The origins of the Internet go back to 1969, to a research project on networking at the Advanced Research Projects Agency of the United States Department of Defense named ARPANET. A number of different network projects subsequently surfaced, with the next landmark developments coming over ten years later. In 1981, BITNET ("Because It's Time") was introduced, providing point-to-point connections between universities for the transfer of electronic mail and files. In 1982, ARPA introduced the Transmission Control Protocol (TCP) and Internet Protocol (IP); TCP/IP allowed different networks to be connected to and communicate with one another, creating the system that is in place today.

Once the machines on a network are connected to one another, there needs to be some way to unambiguously specify a single computer so that messages and files find their intended recipient. To accomplish this, all machines directly connected to the Internet have an **IP address**. IP addresses are unique, identifying one and only one machine. The IP address is made up of four numbers separated by periods; for example, the IP address for the main file server at the National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH) is *130.14.25.1*. The numbers themselves represent, from left to right, the domain (*130.14* for the NIH), the subnet (*.25* for the National Library of Medicine at NIH), and the machine itself (*.1*). While these IP numbers aid the computers in directing data, they are obviously very difficult for users to remember. Therefore, an IP address often has a **fully-qualified domain name** (FQDN) associated with it, which is dynamically translated in the background by a **domain name server**. Going back to the NCBI example, instead of using *130.14.25.1* to access the NCBI computer, a user could instead use *ncbi.nlm.nih.gov* and achieve the same result. Reading from left to right, the IP address goes from least to most specific, while the FQDN equivalent goes from most specific to least. The name of any given computer can then be thought of as taking the general form *computer.domain*, with the top-level domain (the portion coming after the last period in the FQDN) falling into one of the six broad categories shown in Table 2.4.1. Outside the United States, the top-level domain names are replaced with a two-letter code specifying the country where the machine is located (e.g., *.ca* for Canada and *.uk* for the United Kingdom).

CONNECTING TO THE INTERNET

Users who cannot plug their computers directly onto the Internet via an Ethernet, 10BaseT, or similar type of connection have two primary ways of accessing the Net: through **on-line services** or **Internet service providers** (**ISPs**). On-line services—e.g., America Online (AOL), CompuServe, and Prodigy—offer a large number of interactive digital services, including information retrieval, electronic mail (e-mail), bulletin boards,

Contributed by Andreas D. Baxevanis and B.F. Francis Ouellette
Current Protocols in Protein Science (1998) 2.4.1-2.4.18
Copyright © 1998 by John Wiley & Sons, Inc.

BIOMEDICAL LIBRARY

Table 2.4.1 Top-Level Domain Names

Top-level domain name	Significance
Inside U.S.	
.com	Commercial site
.edu	Educational site
.gov	Government site
.mil	Military site
.net	Gateway or network host
.org	Private (usually not-for-profit) organization
Outside U.S. (examples)	
.ca	Canadian site
.ac.uk	Academic site in the United Kingdom
.co.uk	Commercial site in the United Kingdom

and "chat rooms" where users who are on line at the same time can converse with each other about any number of subjects. While the on-line services now provide access to the World Wide Web (see discussion of The World Wide Web), most of the features and services available through these systems reside within a proprietary, closed network; once a connection is made between the user's computer and the on-line service, accessing the special features, or content, of these systems does not require ever leaving the on-line system's host computer. Specialized content can range from access to on-line travel reservation systems to encyclopedias that are constantly being updated—items that would not be available to anyone unless they subscribed to that particular on-line service. The focus on in-house content comes at a price; most of these services impose hourly charges for use, which can add up fairly quickly with even moderate use.

Internet service providers, or ISPs, take the opposite tack. Instead of focusing on providing content, the ISPs provide the tools necessary for users to send and receive e-mail, upload and download files, and navigate around the World Wide Web to find information at remote locations. Major players here include large firms such as AT&T and MCI. However, because it is not necessary for ISPs to provide content, a small cottage industry has quickly arisen, with many small, local firms providing reliable connections to the Internet. The major advantage of ISPs is connection speed; ISPs often provide faster connections than the on-line services. Most ISPs charge a monthly fee for unlimited use.

The distinctions between on-line services and ISPs have begun to blur in the favor of the on-line services. AOL recently changed its pricing structure to a monthly flat fee rather than per-hour charges, so for the same cost charged by many ISPs users can obtain all of the proprietary content found on AOL as well as all of the Internet tools available through ISPs. The extensive AOL network puts access to AOL as close as a local phone call in most of the United States, providing access to e-mail no matter where the user is located—a feature that small, local ISPs cannot match. Developments such as this, coupled with the move of local telephone and cable companies to providing Internet access through new, faster fiber optic networks, foretell major changes in how people will access the Net in the future, changes that should favor the end user both in price and performance.

BCU IACR FREE UNIVERSITY LIBRARY

ELECTRONIC MAIL

Most people are introduced to the Internet through the use of electronic mail (**e-mail**). The use of e-mail has become practically indispensable in many settings because of the convenience that it affords in sending, receiving, and replying to messages. Its advantages include the following:

1. it is much quicker than postal, or "snail mail";
2. it provides a convenient method by which messages can be filed or stored;
3. there is little or no cost involved in sending an e-mail message.

While these and other advantages have pushed e-mail to the forefront of interpersonal communication in both industry and the academic community, there are two major disadvantages that users should be aware of. First is the issue of security. As mail travels towards its recipient, it may pass through a number of remote nodes. The message could be intercepted and read at any one of those nodes by someone with high-level access, such as a systems administrator. Second is the issue of privacy. In industrial settings, e-mail is often considered to be an asset of the company for use only in official communication and, as such, is subject to monitoring by supervisors. The opposite is often true in academic, quasi-academic, or research settings; for example, National Institutes of Health policy encourages personal use of e-mail within the bounds of certain published guidelines. The key words here are "published guidelines"; no matter what the setting, users of e-mail systems should always be informed as to their organization's policy regarding the appropriate use and confidentiality of e-mail so that they may use the tool properly and effectively. An excellent, basic guide to the effective use of e-mail has recently been published and is highly recommended (Lamb and Peek, 1995).

Sending E-Mail

E-mail addresses take the general form *user@computer.domain,* where *user* is the name of the individual user and *computer.domain* specifies the actual computer that the e-mail account is located on. Like a postal letter, an e-mail message is comprised of an **envelope** or **header**, showing the e-mail addresses of the sender and recipient, a line indicating the subject of the e-mail, and information about how the e-mail message actually travelled from the sender to the recipient. The header is followed by the actual message, or **body**, analogous to what would go inside the postal envelope.

E-mail programs vary widely, depending on both the platform and the needs of the users. Usually the characteristics of the local area network (LAN) dictate what types of mail programs can be used, and the decision is often left to systems administrators rather than individual users. Among the most widely used e-mail packages with a graphical user interface are Eudora for the Macintosh and Microsoft Exchange for Macintosh, Windows, and UNIX platforms. Text-based e-mail programs, which are accessed by logging into a UNIX-based account, include Elm and Pine.

Sending E-mail using Pine

Since most Macintosh and IBM-compatible e-mail programs employing graphical user interfaces are fairly simple to use, we will concentrate here instead on Pine, one of the most widely-used UNIX-based mail programs available. To invoke Pine from the UNIX prompt, simply type:

```
unix> pine
```

where *unix>* represents the prompt and the characters in bold following the prompt represent what the user types at the keyboard. Once Pine starts up, the user will see a

BOURNEMOUTH UNIVERSITY LIBRARY

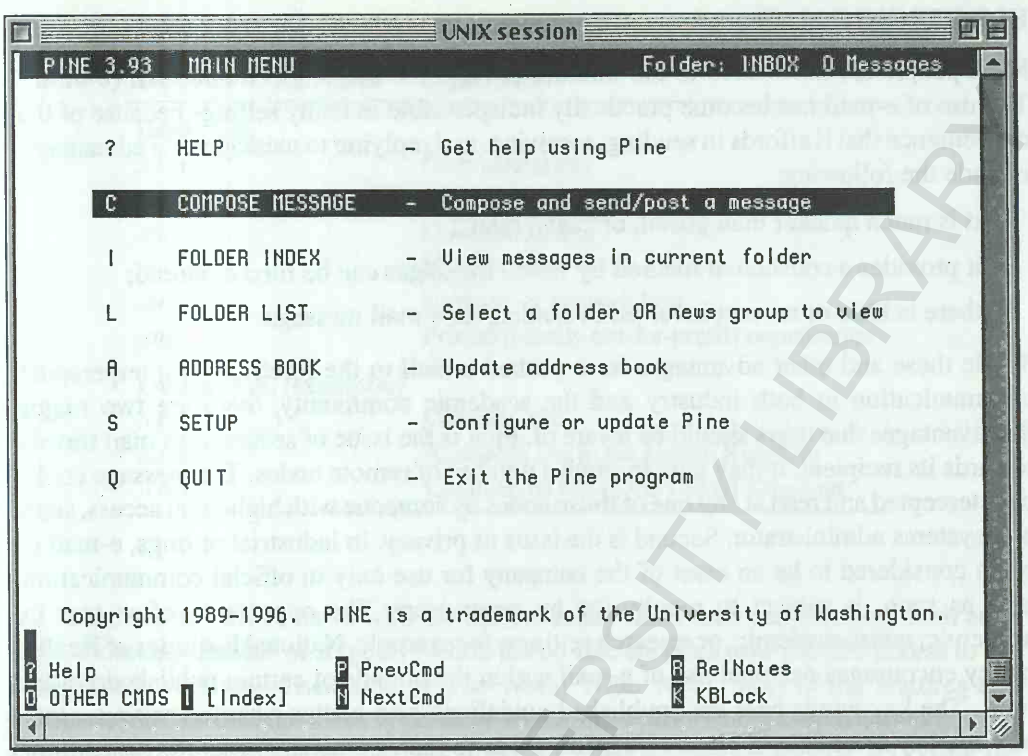

UNIX session
PINE 3.93 MAIN MENU Folder: INBOX 0 Messages
? HELP - Get help using Pine
C COMPOSE MESSAGE - Compose and send/post a message
I FOLDER INDEX - View messages in current folder
L FOLDER LIST - Select a folder OR news group to view
A ADDRESS BOOK - Update address book
S SETUP - Configure or update Pine
Q QUIT - Exit the Pine program
Copyright 1989-1996. PINE is a trademark of the University of Washington.
? Help P PrevCmd R RelNotes
O OTHER CMDS [[Index] N NextCmd K KBLock

Figure 2.4.1 Pine main menu.

menu presenting six major options (Fig. 2.4.1). Any one of these options can be selected either by using the arrow keys and then hitting the Return key or by simply typing in the single-letter command appearing at the left of each line. In order to compose a new e-mail, the user can simply hit the *C* key (for *Compose*) on the keyboard.

Once *C* is selected, a new screen appears, and the user can now fill in the e-mail address of the recipient, the subject of the message, and the actual message itself (Fig. 2.4.2). At the bottom of the screen are a number of commands that apply to this e-mail message, the two most important being *control-C* to cancel the message and *control-X* to send the message. Once either of these is entered, Pine asks for confirmation before sending or deleting the message.

Users can also "attach" files such as word-processing documents or spreadsheets to their e-mail messages. The attached file is sent with the e-mail message. The recipient can then save and use the attached file, provided that they have the appropriate software. To do so, the user would begin composing a message as described above. In order to begin the attachment, one would key in *control-J* to signify that a file is to be attached to the e-mail. At this point, Pine asks for the name of the file to be attached. If the user does not remember the file name, the user can type *control-T* to bring up a list of files from which the desired file can be selected (Fig. 2.4.3).

Files can be selected by moving around the list using the arrow keys on the keyboard. Files have their file sizes next to them, while subdirectories have *dir* listed next to them. When the desired file or directory is highlighted, press *S* for Select. Notice in the example that *parent dir* is highlighted; choosing *parent dir* moves the user one directory up in the file structure. Once a file is selected (rather than a directory), the user is returned to the *Compose* window, where a comment can be added to the e-mail message. At this point, the user can either type the comment and hit *Return*, or just hit *Return* if no comment is

BCU IASH FRONTIER UNIVERSITY LIBRARY

desired. Completion of the composition now resumes as described above. (Since Pine operates in the UNIX environment, the only files that can be attached to an e-mail message are those already on the UNIX machine. Also, even though a file is attached, the recipient may not be able to read the file if they are using a different mail package; consult your systems administrator for more details.)

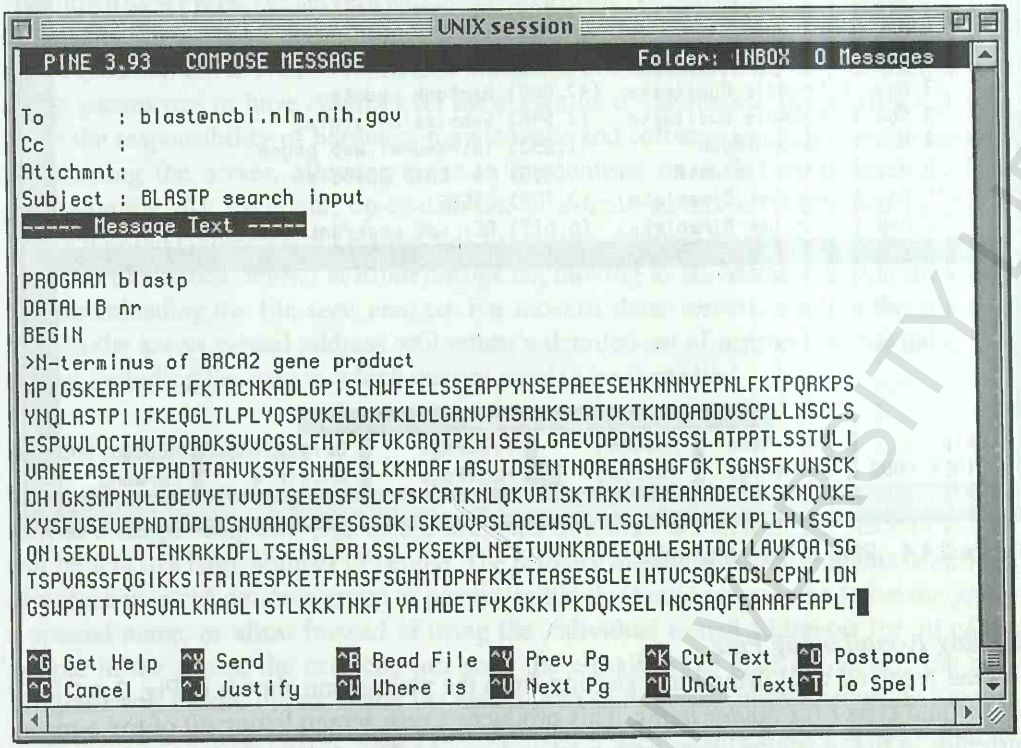


Figure 2.4.2 Pine *Compose* window (with BLAST query as e-mail message).

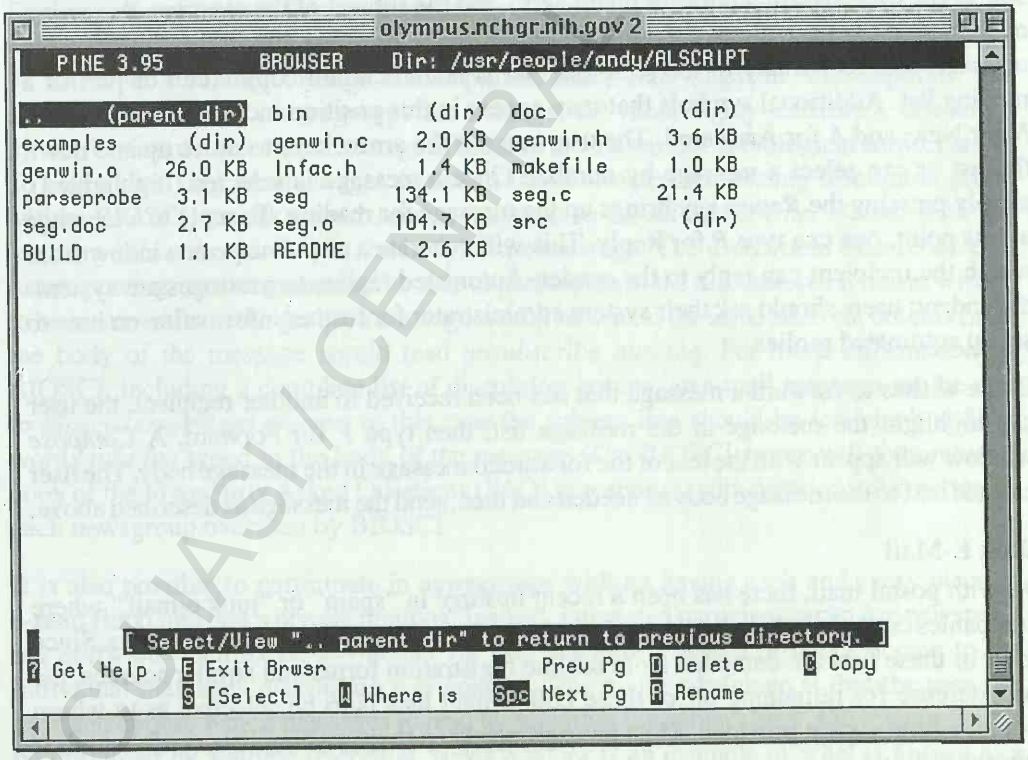

olympus.nchgr.nih.gov 2
PINE 3.95 BROWSER Dir: /usr/people/andy/ALSCRIPT
.. (parent dir) bin (dir) doc (dir)
examples (dir) genwin.c 2.0 KB genwin.h 3.6 KB
genwin.o 26.0 KB lnfac.h 1.4 KB makefile 1.0 KB
parseprobe 3.1 KB seg 134.1 KB seg.c 21.4 KB
seg.doc 2.7 KB seg.o 104.7 KB src (dir)
BUILD 1.1 KB README 2.6 KB
[Select/View ".. parent dir" to return to previous directory.]
? Get Help E Exit Brwsr - Prev Pg D Delete C Copy
S [Select] W Where is Spc Next Pg R Rename

Figure 2.4.3 Pine file browser.

BOLIASI / CENTRAL UNIVERSITY LIBRARY

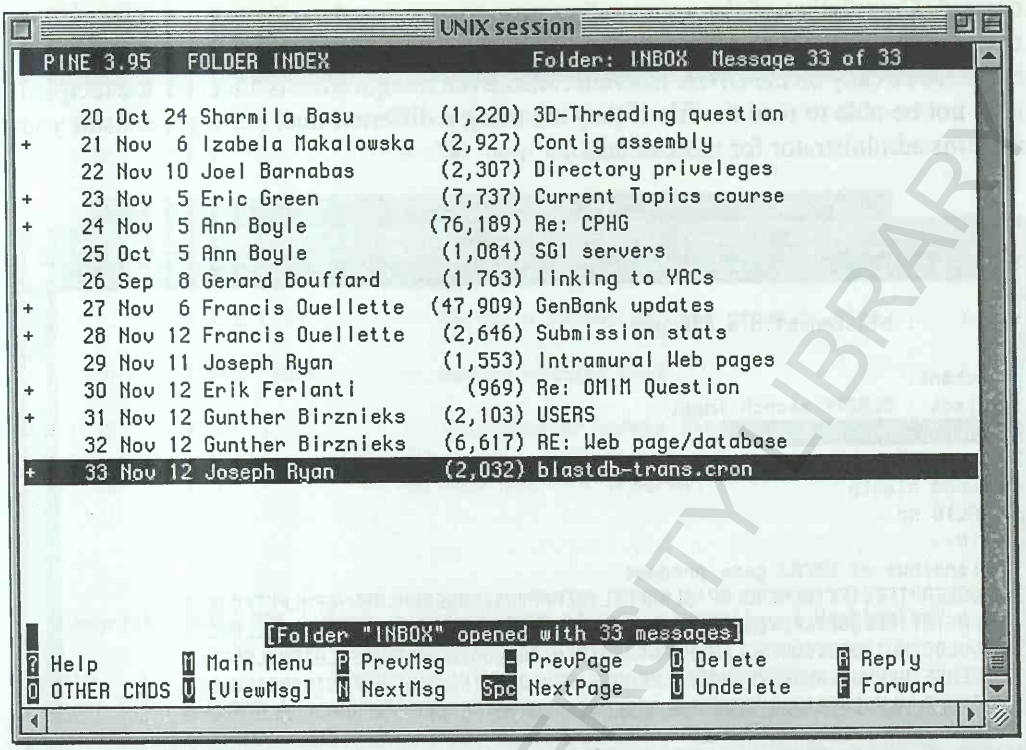

UNIX session
PINE 3.95 FOLDER INDEX Folder: INBOX Message 33 of 33
20 Oct 24 Sharmila Basu (1,220) 3D-Threading question
+ 21 Nov 6 Izabela Makalowska (2,927) Contig assembly
22 Nov 10 Joel Barnabas (2,307) Directory priveleges
+ 23 Nov 5 Eric Green (7,737) Current Topics course
+ 24 Nov 5 Ann Boyle (76,189) Re: CPHG
25 Oct 5 Ann Boyle (1,084) SGI servers
26 Sep 8 Gerard Bouffard (1,763) linking to YACs
+ 27 Nov 6 Francis Ouellette (47,909) GenBank updates
+ 28 Nov 12 Francis Ouellette (2,646) Submission stats
29 Nov 11 Joseph Ryan (1,553) Intramural Web pages
+ 30 Nov 12 Erik Ferlanti (969) Re: OMIM Question
+ 31 Nov 12 Gunther Birznieks (2,103) USERS
32 Nov 12 Gunther Birznieks (6,617) RE: Web page/database
+ 33 Nov 12 Joseph Ryan (2,032) blastdb-trans.cron
[Folder "INBOX" opened with 33 messages]
? Help M Main Menu P PrevMsg - PrevPage D Delete R Reply
O OTHER CMDS V [ViewMsg] N NextMsg Spc NextPage U Undelete F Forward

Figure 2.4.4 Pine e-mail folder index.

Reading E-mail using Pine

To read received mail using Pine, starting from the main menu shown in Fig. 2.4.1 the user would type *I* for *Folder Index*. This produces a new screen listing all of the e-mail currently in the user's mailbox, both read and unread (Fig. 2.4.4). Each message has a message number followed by the date the message was sent, the name of the sender, the size of the message (in number of characters), and the subject of the message. A number of plus signs appear along the left-hand side of the message list, indicating that the message was sent solely to the recipient and is not a "carbon copy" (cc:) or part of a mailing list. Additional symbols that may appear in this position include *D* for Deleted, *N* for New, and *A* for Answered. The user can use the arrow keys to move up and down the list or can select a message by number. Once a message is selected (highlighted), simply pressing the *Return* key brings up the message for reading. To reply to a message at this point, one can type *R* for Reply. This will generate a new *Compose* window from which the recipient can reply to the sender. Automated replies to messages are system-dependent; users should ask their system administrator for further information on how to set up automated replies.

If one wishes to forward a message that has been received to another recipient, the user can highlight the message in the message list, then type *F* for Forward. A *Compose* window will appear with the text of the forwarded message in the message body. The user can add text to the message body as needed and then, send the message as described above.

Bulk E-Mail

As with postal mail, there has been a recent upsurge in "spam" or "junk e-mail," where companies compile bulk lists of e-mail addresses for use in commercial promotions. Since most of these lists are compiled from on-line registration forms and similar sources, the best defense for remaining off of these bulk e-mail lists is to be selective as to whom e-mail addresses are provided. Most newsgroups keep their mailing lists confidential; if in doubt and this is a concern, one should ask.

BCU IASH / CENTRAL UNIVERSITY LIBRARY

E-Mail Servers

So far, the discussion has centered on sending messages, whether it be to one or many recipients. It is also possible to use e-mail as a mechanism for making biological predictions or retrieving records from biological databases. Users can send e-mail messages in a predefined format, defining the action to be performed for remote computers known as **servers**; the servers will then perform the desired operation and e-mail back the results. Figure 2.4.2 actually shows such an e-mail request, using the BLAST e-mail server (see *UNIT 2.5*) at NCBI. While this method is not interactive (in that the user cannot adjust parameters or have control over the execution of the method in real time), it does place the responsibility of hardware maintenance and software upgrades on the persons maintaining the server, allowing users to concentrate on their results instead of on programming. An excellent, up-to-date list of e-mail servers is maintained by Amos Bairoch of the University of Geneva; this list can be obtained by anonymous file transfer protocol (described below) at *expasy.hcuge.ch*, moving to the */databases/info* directory, and downloading the file *serv_ema.txt*. For most of these servers, sending the message *help* to the server e-mail address will return a detailed set of instructions for using that server, including the way in which queries need to be formatted.

Aliases and Newsgroups

In the discussion above (see discussion of Sending E-Mail), the e-mail message is being sent to a single recipient. One of the strengths of e-mail is that a single piece of e-mail can be sent to a large number of people. The primary mechanism for doing this is through **aliases**; a user can define a group of people within their mail program and give the group a special name, or alias. Instead of using the individual e-mail addresses for all of the people in the group, the user can just send the e-mail to the alias name, and the mail program will handle broadcasting the message to each person in that group. Setting up alias names is a tremendous time-saver even for small groups; it also ensures that all members of a given group actually receive all e-mail messages intended for the group.

The second mechanism for broadcasting messages is through **newsgroups**. This model works slightly differently in that the list of e-mail addresses is compiled and maintained on a remote computer through subscriptions, much like magazine subscriptions. For example, the BIOSCI newsgroups are amongst the most highly-trafficked, offering a forum for discussion or the exchange of ideas in a wide variety of biological subject areas. To begin receiving the messages posted to the automated sequencing discussion group within BIOSCI, a user would send a message to *biosci-server@net.bio.net* with the wording *subscribe autoseq* in the body of the message. The user would then receive all future postings to that group and be able to participate in the discussions. If a user wished to be removed from the group, a message would be sent to the same address, but this time, the body of the message would read *unsubscribe autoseq*. For more information on BIOSCI, including a complete list of discussion groups, an e-mail message can be sent to *biosci-server@net.bio.net*; in this case the subject line should be left blank and the words *info faq* typed in the body of the message. The BIOSCI server will then return a copy of the Frequently Asked Questions (FAQ) in response, with detailed information on each newsgroup overseen by BIOSCI.

It is also possible to participate in newsgroups without having each and every piece of e-mail flood into one's private mailbox. Instead, interested participants can use newsreading software, such as NewsWatcher for the Macintosh, which provides access to the individual messages making up a discussion. The major advantage is that the user can pick and choose which messages to read by scanning the subject lines; the remainder can be discarded by a single operation. NewsWatcher is an example of what is known as a

BIOTECHNOLOGY UNIVERSITY LIBRARY

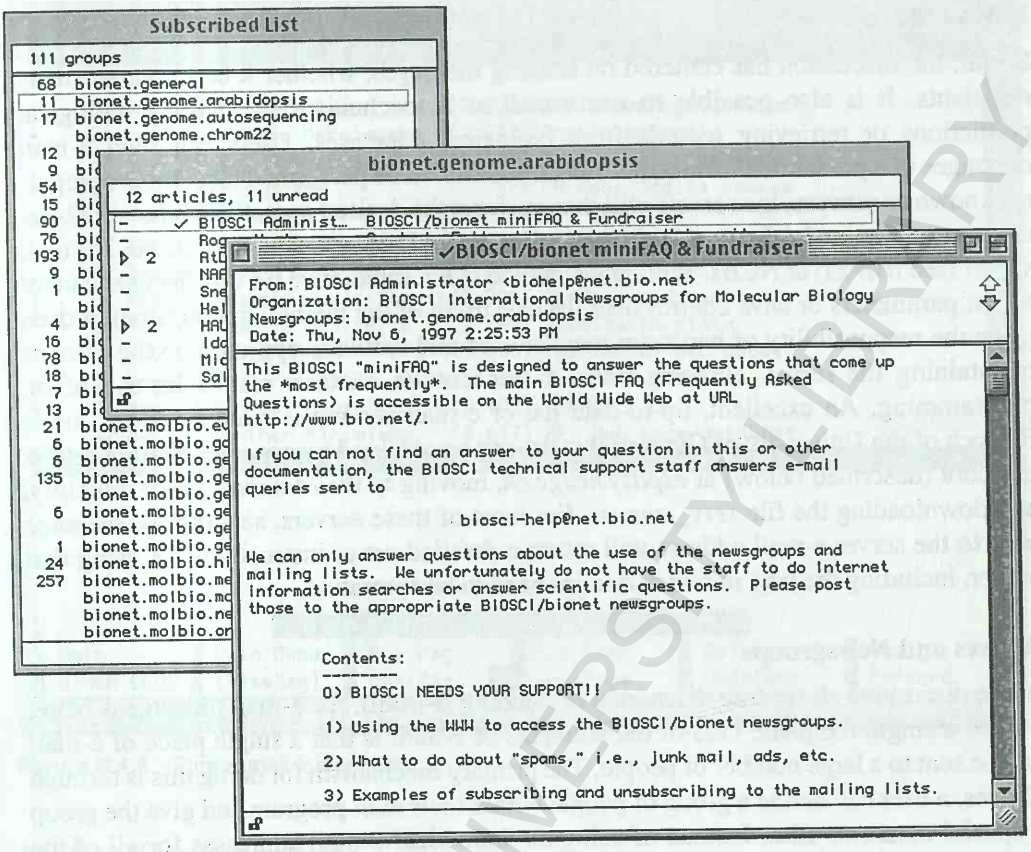

Subscribed List
111 groups
68 bionet.general
11 bionet.genome.arabidopsis
17 bionet.genome.autosequencing
bionet.genome.chrom22
bionet.genome.arabidopsis
12 articles, 11 unread
BIOSCI Administ... BIOSCI/bionet miniFAQ & Fundraiser
✓BIOSCI/bionet miniFAQ & Fundraiser
From: BIOSCI Administrator <biohelp@net.bio.net>
Organization: BIOSCI International Newsgroups for Molecular Biology
Newsgroups: bionet.genome.arabidopsis
Date: Thu, Nov 6, 1997 2:25:53 PM
This BIOSCI "miniFAQ" is designed to answer the questions that come up the *most frequently*. The main BIOSCI FAQ (Frequently Asked Questions) is accessible on the World Wide Web at URL http://www.bio.net/.
If you can not find an answer to your question in this or other documentation, the BIOSCI technical support staff answers e-mail queries sent to
biosci-help@net.bio.net
We can only answer questions about the use of the newsgroups and mailing lists. We unfortunately do not have the staff to do Internet information searches or answer scientific questions. Please post those to the appropriate BIOSCI/bionet newsgroups.
Contents:
0) BIOSCI NEEDS YOUR SUPPORT!!
1) Using the WWW to access the BIOSCI/bionet newsgroups.
2) What to do about "spams," i.e., junk mail, ads, etc.
3) Examples of subscribing and unsubscribing to the mailing lists.

Figure 2.4.5 NewsWatcher windows.

client-server application—the client software (here, NewsWatcher) runs on a client computer (a Macintosh), which in turn interacts with a machine at a remote location (the server). Client-server architecture is interactive in nature, with a direct connection being made between the client and server machines.

Once NewsWatcher is started, the user is presented with a list of newsgroups available to them (Fig. 2.4.5); this list will vary, depending on the user's location, as systems administrators have the discretion to allow or block certain groups at a given site. From the rear-most window in the figure, the user double-clicks on the newsgroup of interest (here, *bionet.genome.arabidopsis*), which spawns the window shown in the center. At the top of the center window is the current unread message count, and any message within the list can be read by double-clicking on that particular line. This, in turn, spawns the last window (in the foreground), showing the actual message. If a user decides not to read any of the messages, or is done reading individual messages, the balance of the messages within the newsgroup (center) window can be deleted by first choosing *Select All* from the *File* menu, then selecting *Mark Read* from the *News* menu. Once the newsgroup window is closed, the unread message count is reset to zero. Every time NewsWatcher is restarted, it will automatically poll the news server for new messages that have been created since the last session.

FILE TRANSFER PROTOCOL

Despite the many advantages afforded by e-mail in transmitting messages, experienced e-mail users have no doubt experienced frustration in trying to transmit files (**attachments**; see discussion of Sending E-Mail) along with an e-mail message. The mere fact that a file can be attached to an e-mail message and sent does not mean that the recipient

BCU IASTCOPY PALUNIVERSITY LIBRARY

will be able to detach, decode, and actually use the attached file. While more cross-platform e-mail packages such as Microsoft Exchange are being developed, the use of different e-mail packages by people at different locations means that sending files via e-mail is not an effective, foolproof method, at least in the short term. One solution to this problem is through the use of a **file transfer protocol** (or **FTP**). The workings of FTP are quite simple—a connection is made between a user's computer (the **client**) and a remote server, and that connection remains in place for the duration of the FTP session. File transfers are very fast, at rates on the order of 5 to 10 kilobytes per second, with speeds varying with time of day, distance between the client and server machines, and overall traffic on the network.

In the ordinary case, making an FTP connection and transferring files requires that a user have an account on the remote server. However, there are many files and programs that the academic community makes freely available, and access to those files does not require having an account on each and every machine where these programs are stored. Instead, connections are made using a system called **anonymous FTP**. Under this system, the user connects to the remote machine, and instead of entering a username/password pair, types *anonymous* as the username and enters an e-mail address in place of a password. Providing one's e-mail address allows the server's systems administrator to compile access statistics which may, in turn, be of use to those actually providing the public files or programs. As with the e-mail servers (see discussion of E-Mail Servers), Amos Bairoch also maintains an extensive list of molecular biology FTP servers. This list can be obtained by connecting to *expasy.hcuge.ch*, moving to the */databases/info* directory, and downloading the file *serv_ftp.txt*.

FTP on a UNIX-based system

For this example, an FTP session will be performed to download the aforementioned list of molecular biology FTP servers. The session begins from the UNIX prompt with the command *ftp*, followed by the name of the remote machine that the user wishes to connect to:

```
ftp expasy.hcuge.ch
```

The bold lettering indicates the characters typed at the keyboard. Once this command is issued, a connect request is made on the remote machine, and the remote machine subsequently asks for the user to log in. As this is an anonymous FTP server, users will type in *anonymous* as their user name, and their e-mail address as the password. The "password" itself will not appear when typing, but is shown below for the sake of example.

```
Connected to expasy.hcuge.ch.
220 expasy FTP server (Version wu-2.4.2-academ[BETA-14](1)
Wed Aug 27 14:44:30 MET DST 1997) ready.
Name (expasy.hcuge.ch:andy): anonymous
331 Guest login ok, send your complete e-mail address as password.
Password: scienceguy1@aol.com
```

Once the connection request is accepted, some introductory information about the FTP site may appear (lines labeled *230*), and eventually an *ftp>* prompt appears.

```
230-Please read the file README
230- it was last modified on Mon Sep 8 18:07:39 1997 - 66 days ago
230 Guest login ok, access restrictions apply.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp>
```

BOWMAN-GITTENAL UNIVERSITY LIBRARY

At this point, the user may start to navigate the directories in the FTP site to find the file of interest. In this case, the user first needs to move to the */databases/info* directory, which is accomplished through a change directory (*cd*) command. The server responds that the "change working directory" (CWD) command was successful. The user may now request the file of interest with the *get* command; the list command (*ls*) can be used to check that the desired file is actually in the current working directory. Once the file transfer is complete, additional files may be downloaded, or the session may be terminated by typing *quit*. The session is terminated when the FTP server issues a *Goodbye*, and the user is returned to the local UNIX prompt.

```
ftp> cd/databases/info
250 CWD command successful.
ftp> get serv_ftp.txt
local: serv_ftp.txt remote: serv_ftp.txt
200 PORT command successful.
150 Opening BINARY mode data connection for serv_ftp.txt (44037 bytes).
226 Transfer complete.
44037 bytes received in 1.29 sec (33.32 Kbytes/s)
ftp> quit
221 Goodbye.
unix>
```

FTP using Fetch

While the above example is not particularly difficult to follow, it does require some basic knowledge and comfort with UNIX commands. Although FTP occurs within the UNIX environment, Macintosh and IBM-compatible computer users can employ programs that utilize graphical user interfaces (GUI, pronounced "gooey") to navigate through the UNIX directories on the FTP server and download files onto their non-UNIX machines. Users need not have any knowledge of UNIX commands to download files; they can instead rely on pop-up menus and the ability to point-and-click their way through the UNIX file structure. The most popular FTP program on the Macintosh platform for FTP sessions is Fetch. To illustrate the difference between UNIX FTP and Fetch, the same file used in the UNIX example above will be downloaded via Fetch on a Macintosh.

Once the Fetch application is launched, a window appears requesting the name of the remote server, the user's name, the user's password, and the target directory (Fig. 2.4.6). This one window combines the first several steps from the UNIX procedure into a single step once the OK button is clicked.

Upon connection, the user is presented with a list of the files appearing in the target directory (Fig. 2.4.7). There is a pull-down menu above the file list; this provides the user a way to "back up" and move out of a folder, back towards the root directory. The file of interest is highlighted by single-clicking its name with the mouse, and the file transfer is initiated when the *Get File* button is pressed. Notice in Figure 2.4.7 that the *Automatic* radio button (near the bottom of the window under the *Get File* button) is selected, meaning that Fetch will determine the appropriate type of file transfer to perform; this may be manually overridden by selecting either *Text* or *Binary*, depending on the nature of the file being transferred. The Macintosh will prompt the user for where the the downloaded file should be stored on the local system. Once a location is chosen, the file is transferred. The connection to the remote server remains in force until the window is closed or the user quits Fetch (using the command *apple-Q*).

In a similar fashion, Fetch can be used to upload files from a Macintosh to a remote computer. In this case, the steps are basically reversed—once the connection to the remote machine is established as described above, the user would simply click the *Put File* button instead. At this point, Fetch will bring up the Macintosh file-dialog window, asking which

BCU IAS CENTRAL UNIVERSITY LIBRARY

file should be uploaded. Once a file is selected, a second window will be spawned, asking how the file should be named once it is transferred to the remote machine as well as how the file should be transferred (most often as either *Text*, *Binary*, or *Raw Data*). As with *Automatic* downloads described in the preceding paragraph, Fetch will attempt to determine the proper protocol based on the type of file being transferred. As a rule, text files should be transferred as *Text*, programs or executables as *Binary*, and graphic format files such as PICT and TIFF files as *Raw Data*.

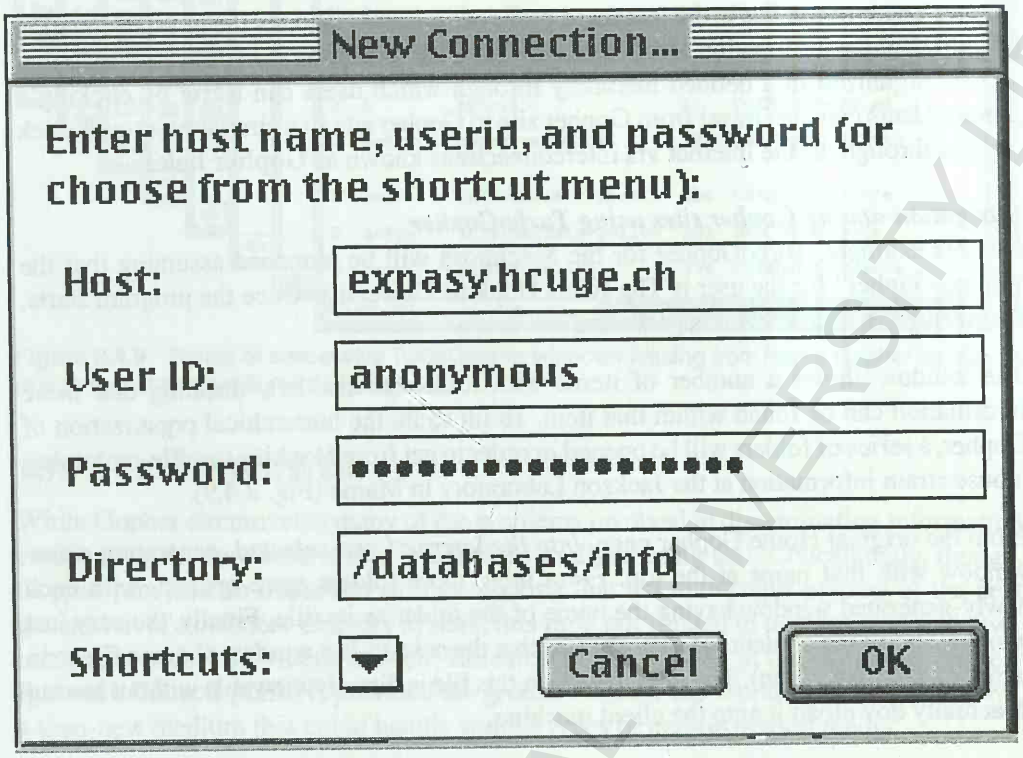


Figure 2.4.6 Fetch window for logging onto host.

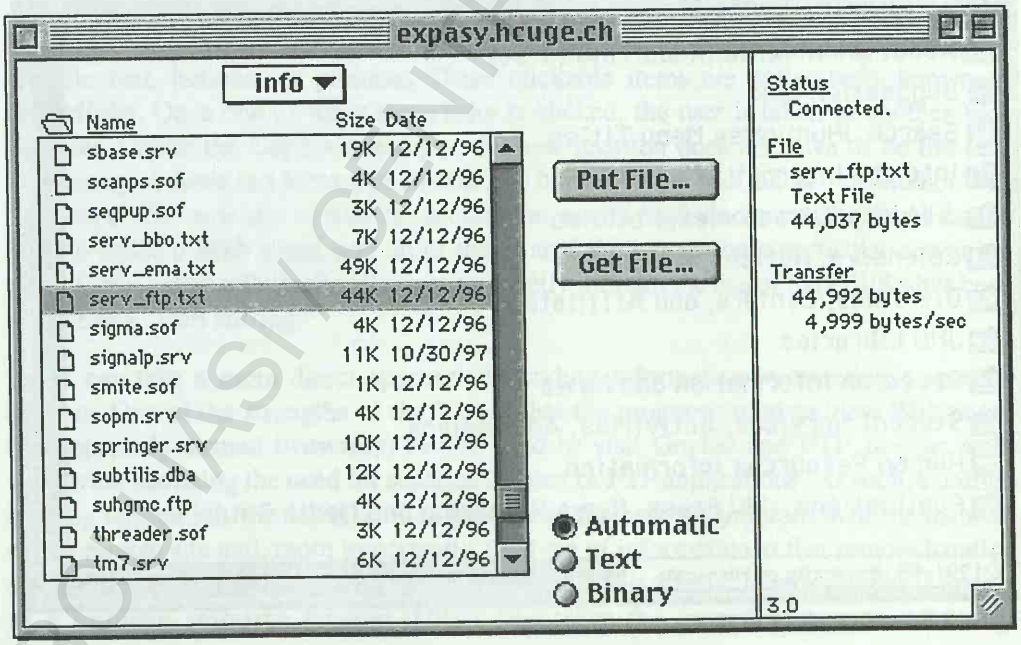


Figure 2.4.7 Fetch window showing files on remote host.

BOSTON CENTRAL UNIVERSITY LIBRARY

GOPHER

When working with FTP, once a user enters a particular directory, they can only see the names of directories or files. In order to actually view what is within those files, it is necessary to physically download files onto one's own computer. To circumvent this problem, researchers at the University of Minnesota developed a system through which text documents can be viewed without actually being downloaded. This program is named **Gopher** in honor of the school's mascot and is one of the first examples of a **distributed document delivery system**, a special type of interactive, client-server application.

Gopher's ease of use lies in the fact that a user does not need to know the physical location or address of the information being sought. All of the information stored at various Gopher sites is organized in a defined hierarchy through which users can travel by clicking a mouse. Users can also travel from Gopher site to Gopher site in a similar point-and-click fashion throughout the Internet via interconnections known as **Gopher holes**.

Navigation among Gopher sites using TurboGopher

For this example, TurboGopher for the Macintosh will be launched assuming that the "Home Gopher" for the user is The Johns Hopkins University. Once the program starts, the window in Figure 2.4.8 is generated.

The window shows a number of items with folders to the left, meaning that more information can be found within that item. To illustrate the hierarchical organization of Gopher, a series of folders will be opened in order to get from Hopkins to a file containing mouse strain information at the Jackson Laboratory in Maine (Fig. 2.4.9).

From the original Home Gopher page, *Into the Internet!* was selected, generating a new window with that name at the top. From there more folders were opened, with each newly-generated window having the name of the folder as its title. Finally, the very last window shows the contents of a file (notice that the next-to-last window shows a file icon rather than a folder icon). The information in this file is directly viewable without having to actually download it onto the client machine.

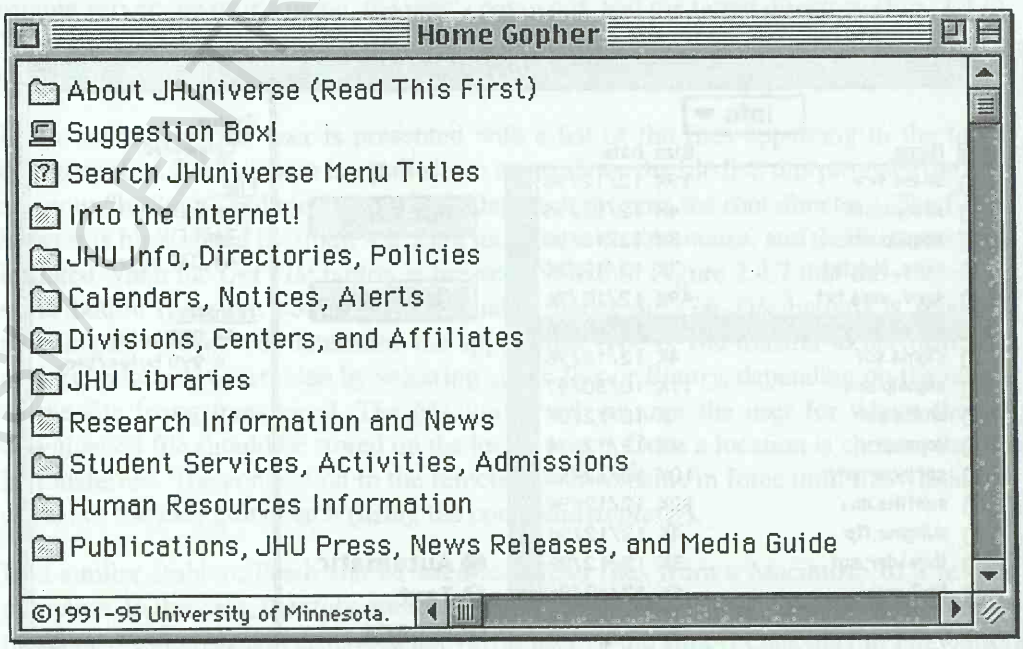


Figure 2.4.8 TurboGopher Home Gopher window.

BCU IASE CENTRAL UNIVERSITY LIBRARY

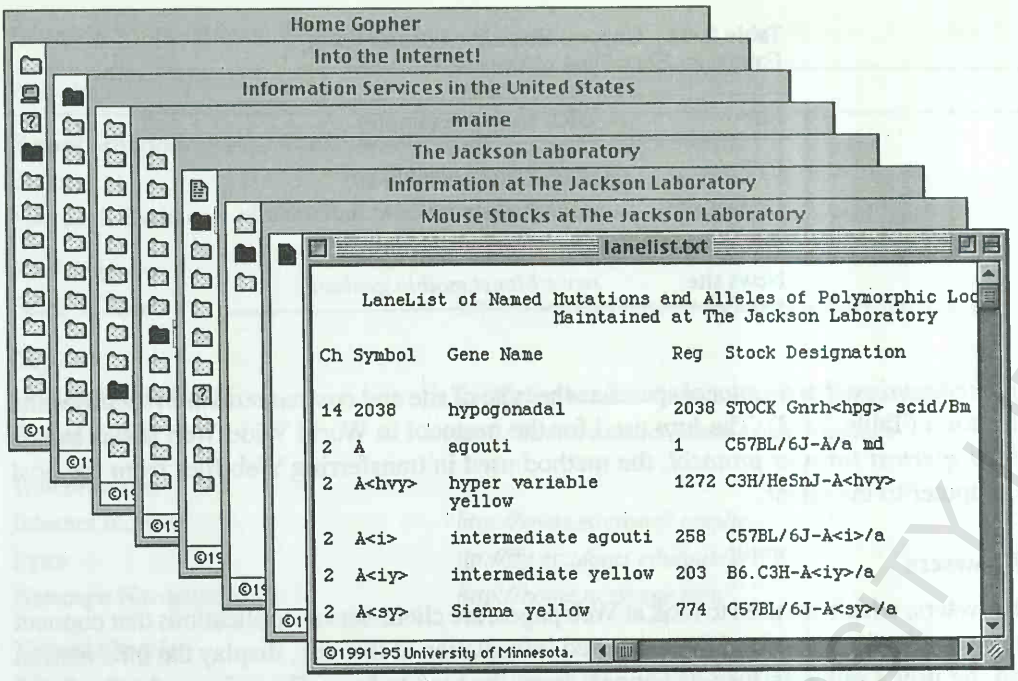

Home Gopher
Into the Internet!
Information Services in the United States
maine
The Jackson Laboratory
Information at The Jackson Laboratory
Mouse Stocks at The Jackson Laboratory
lanelist.txt
LaneList of Named Mutations and Alleles of Polymorphic Loc
Maintained at The Jackson Laboratory
Ch Symbol Gene Name Reg Stock Designation
14 2038 hypogonadal 2038 STOCK Gnrh<hpg> scid/Bm
2 A agouti 1 C57BL/6J-A/a md
2 A<hvy> hyper variable yellow 1272 C3H/HeSnJ-A<hvy>
2 A<i> intermediate agouti 258 C57BL/6J-A<i>/a
2 A<iy> intermediate yellow 203 B6.C3H-A<iy>/a
2 A<sy> Sienna yellow 774 C57BL/6J-A<sy>/a
©1991-95 University of Minnesota.

Figure 2.4.9 Series of successive TurboGopher windows leading from Home Gopher window in rear to text of a target file in the foreground.

THE WORLD WIDE WEB

While Gopher circumvents many of the problems involved in disseminating information easily, its main drawback is that it is limited to distributing text. Conceptually, though, Gopher provided an important shift in thinking that led to the development of the next generation of document-delivery system, this time not limited to text but able to deliver images, sounds, and videos as well. Research begun in 1989 at the European Nuclear Research Council (CERN) provided the groundwork for the World Wide Web (WWW), a then-new medium that could handle various types of non-text-based media.

Navigation on the World Wide Web

As with Gopher, navigation on the Web does not require advance knowledge of the location of the information being sought. Instead, users can navigate by clicking on specific text, buttons, or pictures. These clickable items are collectively known as **hyperlinks**. Once one of these hyperlinks is clicked, the user is taken to another Web location. Unlike the Gopher paradigm, the new location does not have to be the next location up or down in a hierarchy. Instead, the hyperlink can take the user to another Web location at the same site or halfway around the world. Each document displayed on the Web is called a **Web page**, and all of the related Web pages on a particular server are collectively called a **Web site**. Navigation strictly through the use of hyperlinks has been nicknamed "Web surfing."

Users can take a more direct approach to finding information by entering a specific address. One of the strengths of the Web is that the programs used to view Web pages (appropriately termed **browsers**) can be used to visit Gopher and FTP sites as well, somewhat obviating the need for separate Gopher or FTP applications. As such, a unified naming convention was introduced to indicate to the browser program both the location of the remote site and, more importantly, the type of information at that remote location so that the browser could properly display the data. This standard-form address is known as a **uniform resource locator** (**URL**), and takes the general form *protocol://com-*

BUDAPEST CENTRAL EUROPEAN UNIVERSITY LIBRARY

Table 2.4.2 Uniform Resource Locator (URL) Format for Each Type of Transfer Protocol

Site	URL format (example)
FTP site	*ftp://ftp.ncbi.nlm.nih.gov*
Gopher site	*gopher://gopher.iubio.indiana.edu*
Web site	*http://www.nhgri.nih.gov*
News site	*news:bionet.molbio.genbank*

puter.domain, where *protocol* specifies the type of site and *computer.domain* specifies the location (Table 2.4.2). The *http* used for the protocol in World Wide Web URLs stands for *hypertext transfer protocol*, the method used in transferring Web files from the host computer to the client.

Browsers

Browsers, which are used to look at Web pages, are client-server applications that connect to a remote site, download the requested information at that site, display the information on the user's monitor, then disconnect from the remote host. The information retrieved from the remote host is in a platform-independent format called **hypertext markup language (HTML)**. HTML code is strictly text-based, and any associated graphics or sound for that document exist as separate files in a common format. For example, images may be stored and transferred in GIF format, developed by CompuServe for the quick and efficient transfer of graphics; while GIF format is most commonly used for graphics, other formats such as JPEG and BMP may also be used. Because of this, a browser can display any Web page on any type of computer, whether it be a Macintosh, IBM-compatible, or UNIX machine. The text is displayed first, then the remaining elements are placed on the page as they are downloaded. With minor exceptions, a given Web page will look the same when the same browser is used on any of the above platforms.

One of the two major browsers is Netscape Communicator. Netscape, developed by Marc Andressen and colleagues, is the outgrowth of the first widely-used Web browser, NCSA Telnet (named for the National Center for Supercomputing Applications). Netscape was able to capture its large market share by basically giving its browser away for years, relying for revenue on Web servers and other products. While Netscape can still be freely downloaded from the Netscape Web site (see Table 2.4.3), newly-downloaded copies can be used free of charge only for the first 90 days (the fee is waived for educational institutions). Netscape does have some notable drawbacks, such as a clumsy bookmark system for storing the addresses of frequently visited sites, difficulty in displaying and navigating Web pages where the screens are split into separate frames, and a tendency to stall or crash frequently. However, it does a more than adequate job in addressing the Web-browser needs of most users.

Netscape's main competition comes from Microsoft, which is aggressively seeking to displace Netscape as the market leader with its own browser, Internet Explorer. Microsoft's strategy is to keep Internet Explorer a freeware product that can be downloaded directly from the Microsoft Web site (see Table 2.4.3). Internet Explorer has some key advantages over Netscape. Pages tend to load faster, the browser is more stable, and it handles frames much better than does Netscape. As the version numbers of both of these browsers advance, users will find many of the same features in both programs, and the final decision by users will come down to a matter of cost and personal preference.

BCU IASTCHAL UNIVERSITY LIBRARY

Table 2.4.3 World Wide Web Sites of Interest

Web site	URL
E-mail programs	
Eudora	*http://www.eudora.com*
Microsoft Exchange	*http://www.microsoft.com/exchange/eval.html*
FTP program	
Fetch 3.0/Mac	*http://www.dartmouth.edu/pages/softdev/fetch.html*
Gopher	
Gopher User's Guide	*ftp://boombox.micro.umn.edu/pub/gopher/docs*
TurboGopher	*http://boombox.micro.umn.edu/hh/gopher/*
Web browsers	
Internet Explorer	*http://www.microsoft.com/ie*
Lynx	*ftp://ftp.cc.ukans.edu/pub/WWW*
Netscape Navigator	*http://home.netscape.com*
Virtual libraries	
EBI BioCatalog	*http://www.ebi.ac.uk/biocat/biocat.html*
Molecular Biology E-Mail servers	*ftp://expasy.hcuge.ch/databases/info/serv_ema.txt*
Molecular Biology FTP servers	*ftp://expasy.hcuge.ch/databases/info/serv_ftp.txt*
Pedro's Biomolecular Research Tools	*http://www.public.iastate.edu/~pedro/research_tools.html*
WWW Virtual Library	*http://www.golgi.harvard.edu/biopages/all.html*
Web search engines	
MetaCrawler	*http://www.metacrawler.com*
SavvySearch	*http://guaraldi.cs.colostate.edu:2000/form*

It is worth mentioning that, while the Web is by definition a visually based medium, it is also possible to travel through Web space and view documents without the associated graphics. For users limited to line-by-line terminals, a browser called Lynx is available. Developed at the University of Kansas, Lynx allows users to use their keyboard arrow keys to highlight and select hyperlinks, using their return key the same way that Netscape and Internet Explorer users would click their mouse. Lynx is available for both the DOS and UNIX platforms (Table 2.4.3).

Internet versus Intranet

The World Wide Web is normally thought of as a way to communicate with people at a distance, but the same infrastructure can be used to connect people within an organization. Such **intranets** provide an easily accessible repository of relevant information, capitalizing on the simplicity of the Web interface. It also provides another channel for broadcast or confidential communication within the organization. Having an intranet is of particular value when members of an organization are physically separated, whether it be in different buildings or different cities. Intranets are protected in such a way that people who are not on the organization's network are prohibited from accessing the internal Web pages; additional protections through the use of passwords are also common.

BCU MAIN CENTRAL UNIVERSITY LIBRARY

Finding Information on the World Wide Web

Most people find information on the Web the old fashioned way—by word of mouth. Alternatively it is possible to find Web sites of interest either by using tables such as Table 2.4.3 or by simply following hyperlinks put in place by Web authors. Continuously clicking from page to page can be a highly ineffective way of finding information, especially when the information sought is of a very focused nature. One way of finding interesting and relevant Web sites is to consult **virtual libraries**, which are curated lists of Web resources arranged by subject. Virtual libraries of special interest to biologists include the WWW Virtual Library, maintained by Keith Robison at Harvard; Pedro's Biomolecular Research Tools, compiled by Pedro Coutinho at Iowa State; and the EBI BioCatalog, based at the European Bioinformatics Institutes. The URLs for these sites can be found in Table 2.4.3.

It is also possible to directly search the Web by using **search engines** such as Alta Vista and Excite, among others. A search engine is simply a specialized program that can perform full-text or keyword searches on databases that catalog Web content. The result of a search is a hyperlinked list of Web sites fitting the search criteria from which the user can visit any or all of the found sites. However, search engines use slightly different methods in compiling their databases. One variation is the attempt to visit and read data from all Web sites ("Web crawling"). Another technique is to catalog only the title of each Web page rather than its entire text. A third is to consider words that must appear next to each other or only relatively close to one another. Because of these differences in search-engine algorithms, the results returned by issuing the same query to a number of different search engines can produce wildly different results. Unless a particular search engine ranks its results by relevance (for example, by scoring words in a title higher than words in the body of the Web page), the results obtained may not be particularly useful.

To address this point, a new class of search engines called **meta-search engines** have been developed that take a more intelligent approach to Web discovery. These programs will take the user's query and poll anywhere from five to ten of the "traditional" search engines. The meta-search engine will then collect the results, filter out duplicates, and return a single, annotated list to the user. One big advantage is that the meta-search engines take relevance statistics into account, returning much smaller lists of results. Even though the hit list is substantially smaller, it is much more likely to contain sites that directly address the original query. Since the programs must poll a number of different search engines, searches conducted this way obviously take longer to perform, but the higher degree of confidence in the compiled results for a given query outweighs the extra few minutes (sometimes only seconds) of search time. Reliable and easy-to-use meta-search engines include SavvySearch and MetaCrawler (see Table 2.4.3).

As an example, let us consider the case where information on the mapping of human chromosome 7 is needed, but no Web address is at hand. In order to find such information without having to visit multiple search sites, a query will be issued through SavvySearch (*http://guaraldi.cs.colostate.edu:2000/form*). The Web form (Fig. 2.4.10) is rather simple to use, with a single box in which search terms can be typed. Here, the search terms *Chromosome 7 Mapping* have been entered; search results are better focused when more words of a specific nature are used. A variety of information sources can be searched, and a number of them have been checked off in the example. If no boxes are checked, SavvySearch defaults to checking Web resources only. Under query options, the most important choice to be made is in the pull-down menu labelled *Search for documents containing*. If *all query terms* is selected, all of the words must appear in the target document, but not in any given order or proximity. If *all query terms, as a phrase* is selected, the words must all appear, and they must all appear together in the order they

BCU IAS CENTRAL UNIVERSITY LIBRARY

were input in the query box. Finally, if *any query terms* is selected, only one of the words need appear in a target document in order for it to appear on the hit list; in practice, this option should never be selected, since it will bring back a large number of pages not really fitting the original query. By checking *Integrate results*, if the same hit is found by multiple search engines, it will only appear once on the final hit list presented to the user.

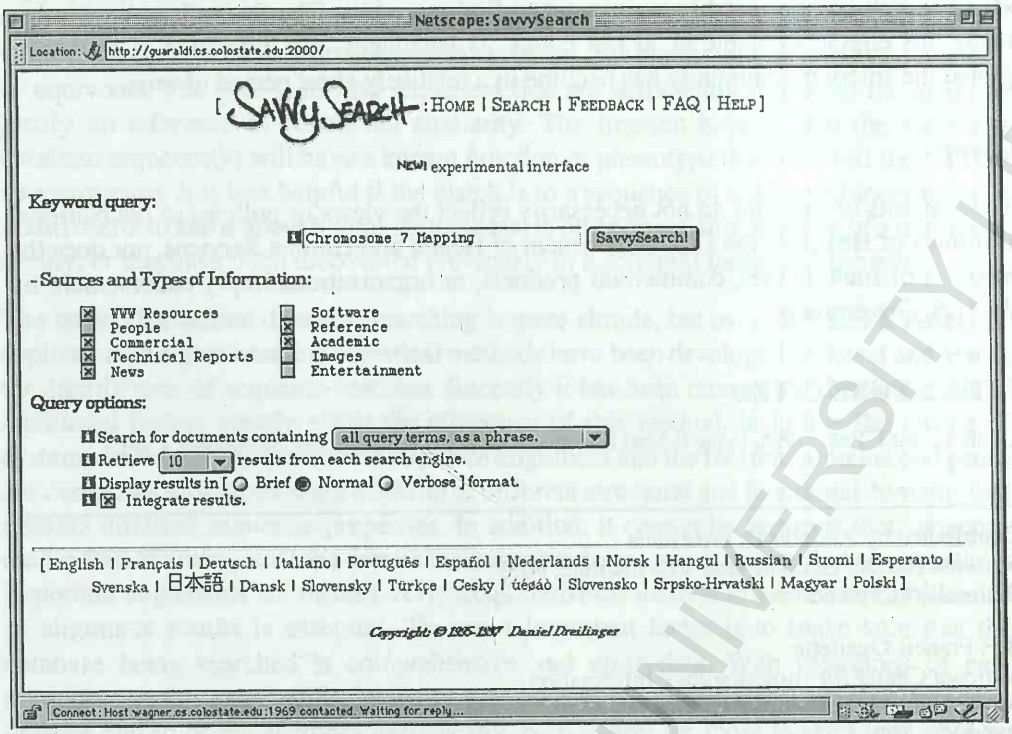

Netscape: SavvySearch
Location: http://guaraldi.cs.colostate.edu:2000/
[SavvySearch: Home | Search | Feedback | FAQ | Help]
NEW! experimental interface
Keyword query:
Chromosome 7 Mapping
SavvySearch!
- Sources and Types of Information:
WWW Resources
People
Commercial
Technical Reports
News
Software
Reference
Academic
Images
Entertainment
Query options:
Search for documents containing all query terms, as a phrase.
Retrieve 10 results from each search engine.
Display results in [Brief Normal Verbose] format.
Integrate results.
[English | Français | Deutsch | Italiano | Português | Español | Nederlands | Norsk | Hangul | Russian | Suomi | Esperanto | Svenska | 日本語 | Dansk | Slovensky | Türkçe | Cesky | ùéøáò | Slovensko | Srpsko-Hrvatski | Magyar | Polski]
Copyright © 1995-1997 Daniel Dreilinger
Connect: Host wagner.cs.colostate.edu:1969 contacted. Waiting for reply...

Figure 2.4.10 SavvySearch query form.

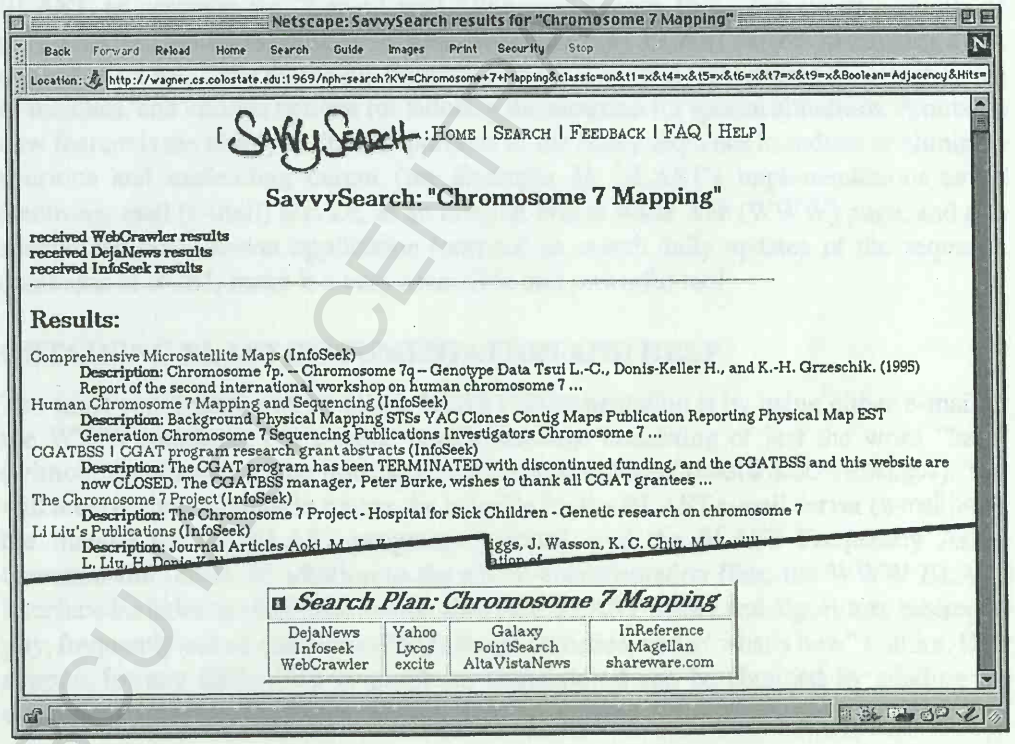

Netscape: SavvySearch results for "Chromosome 7 Mapping"
Back Forward Reload Home Search Guide Images Print Security Stop
Location: http://wagner.cs.colostate.edu:1969/nph-search?KW=Chromosome+7+Mapping&classic=on&t1=x&t4=x&t5=x&t6=x&t7=x&t9=x&Boolean=Adjacency&Hits=
[SavvySearch: Home | Search | Feedback | FAQ | Help]
SavvySearch: "Chromosome 7 Mapping"
received WebCrawler results
received DejaNews results
received InfoSeek results
Results:
Comprehensive Microsatellite Maps (InfoSeek)
Description: Chromosome 7p. -- Chromosome 7q -- Genotype Data Tsui L.-C., Donis-Keller H., and K.-H. Grzeschik. (1995) Report of the second international workshop on human chromosome 7 ...
Human Chromosome 7 Mapping and Sequencing (InfoSeek)
Description: Background Physical Mapping STSs YAC Clones Contig Maps Publication Reporting Physical Map EST Generation Chromosome 7 Sequencing Publications Investigators Chromosome 7 ...
CGATBSS | CGAT program research grant abstracts (InfoSeek)
Description: The CGAT program has been TERMINATED with discontinued funding, and the CGATBSS and this website are now CLOSED. The CGATBSS manager, Peter Burke, wishes to thank all CGAT grantees ...
The Chromosome 7 Project (InfoSeek)
Description: The Chromosome 7 Project - Hospital for Sick Children - Genetic research on chromosome 7
Li Liu's Publications (InfoSeek)
Description: Journal Articles Aoki, M., L. W... Biggs, J. Wasson, K. C. Chiu, M. Yawitt, L. Liu, H. Donis-K... billon of
Search Plan: Chromosome 7 Mapping
DejaNews Infoseek WebCrawler
Yahoo Lycos excite
Galaxy PointSearch AltaVistaNews
InReference Magellan shareware.com

Figure 2.4.11 SavvySearch hit list.

DCU CENTRAL UNIVERSITY LIBRARY

The resultant hit list (Fig. 2.4.11) shows which search engines were used to produce the final hit list, as well as a list of documents fitting the search query as issued. For each hit, a hyperlink is provided that the user can click to go directly to that page. Also shown are the first few lines of the Web page, so that the user can assess whether that particular page is of use to them prior to actually loading the page. At the bottom of the page appears a box offering alternative *Search Plans*, which can be used to cast the net beyond the original set of search engines used in an attempt to find more hits. The SavvySearch team has made this engine available in, at last count, 23 languages—an indication of how truly global the Internet community has become in a relatively short period of time.

NOTE

The contents of this unit do not necessarily reflect the views or policies of the National Institutes of Health or the U.S. Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

LITERATURE CITED

Lamb, L. and J. Peek. 1995. Using E-Mail Effectively. O'Reilly & Associates, Sebastopol, Calif.

Contributed by Andreas D. Baxevanis
National Human Genome Research Institute, NIH
Bethesda, Maryland

B.F. Francis Ouellette
National Center for Biotechnology Information
National Library of Medicine, NIH
Bethesda, Maryland

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Sequence Similarity Searching Using the BLAST Family of Programs

Database "homology" searching is carried out thousands of times each day by scientists worldwide; it is a critical tool for all molecular biologists. The basic operation is to take a query sequence (consisting of either nucleotides or amino acids), compare it sequentially with all known sequences in a database, and rank the database sequences in a "hit list" with respect to how closely they resemble the query. The results may be positive, negative, or equivocal. The goal is to find sequences that are sufficiently similar to the query to justify an inference of functional similarity. The implicit hope is that the matching database sequence(s) will have a known function or phenotype that will shed light on the new sequence. It is less helpful if the match is to a sequence of unknown function, but it is still useful to know about homologs that exist in other organisms or as part of a multigene family, or that have been discovered previously in a different biological context.

The basic idea behind database searching is quite simple, but over the years a variety of sophisticated algorithms and statistical methods have been developed to detect and assess the significance of sequence matches. Recently it has been recognized that a number of additional factors greatly affect the efficiency of this method, including the choice of systems for assigning a score to a sequence alignment and the fact that proteins and genes are composite structures often made up of different structural and functional domains that possess different sequence properties. In addition, it cannot be assumed that sequence data is free of errors, and thus certain measures are necessary to ensure that no potentially important alignments are missed. A rigorous statistical method to assess the significance of alignment results is essential. The most important factor is to make sure that the database being searched is comprehensive and up-to-date. With thousands of new sequences entering the public databases every week, it is essential to have access to daily updates and to repeat searches periodically, particularly for those queries that yield no significant matches upon initial analysis.

BLAST, an acronym for "Basic Local Alignment Search Tool," represents a family of programs for database similarity searching (Table 2.5.1). BLAST programs embody a fast algorithm for sequence matching, rigorous statistical methods for judging the significance of matches, and various options for tailoring the program for special situations. A notable new feature is the ability to "mask" portions of the query sequence to reduce or eliminate spurious and misleading output (see Example 3). BLAST's implementations as an electronic mail (e-mail) service, as an Internet World Wide Web (WWW) page, and as a specialized client-server application (*UNIT 2.4*) to search daily updates of the sequence databases at NCBI, make it a very accessible and powerful tool.

OBTAINING BLAST DOCUMENTATION AND HELP

The most convenient way to access BLAST documentation is by using either e-mail or the WWW interface. For e-mail, send a message consisting of just the word "help" (without the quotes) to the normal BLAST server address (blast@ncbi.nlm.nih.gov). You will receive three e-mails in return: the help file for the BLAST e-mail server (e-mail help), the manual for the BLAST programs (manual), and the BLAST Frequently Asked Question file (FAQ). In addition to the above documentation files, the WWW BLAST interface includes on-line information about the BLAST theory and algorithm, bibliography, frequently-asked questions (FAQ), usage statistics, and a "what's new" feature. User support for any difficulties or problems encountered can be obtained by sending the question to blast-help@ncbi.nlm.nih.gov.

Contributed by Hervé E. Recipon, Gregory D. Schuler, Mark S. Boguski, and Warren Gish
Current Protocols in Protein Science (1996) 2.5.1-2.5.38
Copyright © 1996 by John Wiley & Sons, Inc.

BOSTON COLLEGE UNIVERSITY LIBRARY

Table 2.5.1 BLAST Family of Programs for Database Similarity Searching

Program	Query sequence	Database sequences	Comments	Example
BLASTN	Nucleotide (both strands)	Nucleotide	Parameters optimized for speed, not sensitivity; not intended for finding distantly related coding sequences; automatically checks complementary strand of query	1,4
BLASTX	Nucleotide (six-frame translations)	Protein	Very useful for preliminary data containing potential frameshift errors, e.g., ESTs and other "single pass" sequences; 14 different genetic codes available; 65 different scoring matrices available; low-complexity filter options with SEG or XNU algorithms (highly recommended)	2,6
BLASTP	Protein	Protein	65 different scoring matrices available; low-complexity filter options	3
TBLASTN	Protein	Nucleotide (six-frame translations)	Essential for searching protein queries against dbEST; often useful for finding undocumented open reading frames or frameshift errors in database sequences; 14 different genetic codes available; 65 different scoring matrices available; low-complexity filter options with SEG and XNU algorithms (highly recommended)	5
TBLASTX	Nucleotide (six-frame translations)	Nucleotide (six-frame translations)	Restricted for search against dbEST, dbSTS, and Alu (except for e-mail server); 14 different genetic codes available; 65 different scoring matrices available; low-complexity filter options with SEG and XNU algorithms (highly recommended)	7

SUBMITTING SEQUENCES FOR DATABASE HOMOLOGY SEARCHING

Making use of remote BLAST services over the Internet (*UNIT 2.4*) offers several distinct advantages over performing searches on a local computer system. These include access to daily comprehensive updates of the sequence databases as well as to sophisticated software to run searches on computers that are much faster and more powerful than those generally available locally.

As outlined below, there are three main ways to submit a BLAST query for remote execution. Regardless of which method is used, there are four basic elements common to all BLAST requests:

1. The BLAST command to be executed;
2. The name of the database to be searched;
3. Additional parameters and modifiers (optional); and
4. The query sequence.

In a short period of time, usually measured in minutes, you will receive the results (BLAST output) of your search. Then you must evaluate whether or not anything interesting was found. There are two basic steps to this process.

BJU IASRI CENTRAL UNIVERSITY LIBRARY

First, determine if any matches are statistically significant. Keep in mind that for any query sequence, all database search programs will produce an ordered list of aligned sequences, but none of them need have any biological significance. The question is what degree of similarity is necessary to be considered surprising—i.e., not likely to have occurred by chance. For each database match, BLAST produces a *p* value derived from the application of a rigorous theory for the assessment of sequence similarities (see Commentary for a detailed discussion of this topic).

Second, assuming some significant or potentially significant matches are found, retrieve the full database records for the sequences of interest to examine their annotation, which often has more extensive information on function, expression, and other topics of interest. The database record also provides links to the scientific literature. The basic procedure is to use the sequence accession number as a key to obtain the database record and associated literature using either e-mail or Internet client-server resources (*UNIT 2.4*). NCBI's RETRIEVE server (accessible by e-mail), World Wide Web/Mosaic page, and Entrez system (available over the Internet, on CD-ROM, and is also included on *Current Protocol on CD-ROM*) are specific examples.

BLAST searches are available via e-mail and a WWW interface and as an Internet service using a command-line interface. All three applications are described in more detail below. E-mail is ubiquitous, convenient, and well-suited to those who only submit occasional queries. The WWW interface makes it easy to "copy and paste" sequences for searching. The command-line interface is the most powerful way to use BLAST and is easy to incorporate into automated systems for high-throughput analysis.

Submitting a Sequence to BLAST Using the Electronic Mail Server

BASIC PROTOCOL 1

Materials

Any computer system and a means of composing, sending, and receiving e-mail
BLAST e-mail server help file (see *UNIT 2.4*; optional but highly recommended)
Query sequence file(s) in FASTA/Pearson format (see Example 1)

NOTE: Most commercial sequence analysis software packages provide file interconversion routines that produce FASTA/Pearson format. A free, public-domain program called READSEQ can interconvert more than a dozen different formats, including FASTA/Pearson. READSEQ, written by Don Gilbert, is available via Gopher or anonymous FTP (see *UNIT 2.4*) from ftp.bio.indiana.edu in the /molbio/readseq directory. Note that the BLAST programs are not case-sensitive with regard to query sequences, which therefore may consist of either upper- or lower-case letters.

1. Create new mail and address it to blast@ncbi.nlm.nih.gov. It is not necessary to supply a subject line, but if one is provided, the server will use it in the subject of its reply.
2. In the body of the message, type the keyword PROGRAM followed by the name of the BLAST command to be executed (BLASTN, BLASTP, BLASTX, TBLASTN, and TBLASTX are available; see Table 2.5.1 and the BLAST e-mail help file for more details).
3. On the next line, type the keyword DATALIB followed by the name of the desired database. Although we use genbank in Example 1 for simplicity, we recommend using nr, which is a comprehensive, quasi-nonredundant superset of GenBank plus nucleotide sequence data from other sources.

BCM/SJCH CENTRAL INVENTORY LIBRARY

4. Start a new line and type the keyword BEGIN to indicate the beginning of your query sequence. Insert the contents of a FASTA-formatted sequence file into the mail message beginning on the next line.
5. Insert a blank line to indicate the end of the sequence (see the FAQ file, questions 17 and 29, for more details).
6. Send the message. You will receive the results by return e-mail, usually within 1 hr. Occasionally it will take longer than this, due to heavy loads on the system or server maintenance. If you do not receive a reply by the next morning, contact NCBI at the BLAST help address, blast-help@ncbi.nlm.nih.gov, or by telephone at 301-496-2475.

Note that many other parameters may be applied to BLAST searches. Some of these are described below. A complete description of e-mail server directives and their uses is included in e-mail help, which also includes descriptions of available databases and an explanation of sequence identifiers.

BASIC PROTOCOL 2

Submitting a Sequence to BLAST Using the Command-Line Client Program

Materials

UNIX computer with TCP/IP and BSD socket support, Ethernet hardware, Internet connection, and C compiler
Suite of BLAST programs appropriate for the computer system
Query sequence file(s) in FASTA format (see Example 1)

NOTE: Compatible operating system support for TCP/IP and BSD sockets is provided by Sun Microsystems, Silicon Graphics, Hewlett-Packard, Digital Equipment Corporation, and IBM. Many modern UNIX workstations come with Ethernet hardware built in. If these terms are unfamiliar to you, please consult your local systems manager for more information.

NOTE: The source code for BLAST programs is available at NCBI via anonymous FTP (ftp to ncbi.nlm.nih.gov) beneath the BLAST directory.

1. At the console command prompt, type the following items in order, separated by spaces:
   a. The BLAST command (e.g., blastn);
   b. The name of the database (e.g., genbank); and
   c. The name of a file containing your query sequence.
2. If desired, change default search parameters or enter optional parameters after the BLAST command on the same line, separated by spaces. A complete list of parameters and options, their syntax and uses are described in the BLAST manual file.
3. It is normally desirable to capture the output in a disk file (by default, it would only be displayed on the console). To do this under UNIX, include in your command the redirection symbol > followed by the name of a file.
4. Press the Enter key to execute the command.

BRU IAS/CENTRAL UNIVERSITY LIBRARY

Submitting a Sequence Using NetBLAST Graphical Client Program

BASIC PROTOCOL 3

Materials

One of the following computer systems equipped with a network interface and having the ability to access the Internet (*UNIT 2.4*):
- Macintosh computer running System 7.0 (or later) operating system and MacTCP system extension
- IBM-compatible computer running Microsoft Windows 3.1 (or later) and any network software package that supports the WinSock specification

NetBLAST–executable program appropriate for the system
Query sequence file(s) in FASTA format (see Example 1)

NOTE: The NetBLAST executable program will be included on *Current Protocols on CD-ROM*. Contact NCBI at net-info@ncbi.nlm.nih.gov or 301-496-2475 regarding current availability.

1. Start the NetBLAST program (see Example 4).
2. As the program begins, a File-Open dialog box will appear. Enter or select the name of a file containing the query sequence. After selecting the query sequence, press the OK button to continue.
3. Another dialog box will appear listing options to further define the search conditions. The settings are organized into categories, represented by "folder tabs" across the top of the window. Click on the Database tab.
4. Select the database type (nucleotide or protein) that you want to search. If the query is a nucleotide sequence, you may check the Translate query option to search a protein database.
5. A list of databases of the given type will be displayed; select the one you want to search.
6. Select other tabs to examine or modify additional parameters by clicking on them (optional).
7. Press the Search button to send the request to the server. When the results are returned, they will be displayed in a window.

Submitting a Sequence Using the WWW Interface

BASIC PROTOCOL 4

Materials

One of the following computer systems equipped with a network interface and the ability to access the Internet (see *UNIT 2.4*):
- UNIX computer with TCP/IP and BSD socket support and Ethernet hardware
- Macintosh computer running the System 7.0 (or later) operating system and the MacTCP system extension
- IBM-compatible computer running Microsoft Windows 3.1 (or later) and any network software package that supports WinSock

WWW browser–executable program appropriate for the system
Query sequence file(s) in FASTA format (see Example 1)

NOTE: Browsers may be obtained at one of the following FTP sites: ftp.ncsa.uiuc.edu for NSCA (UNIX, Mac, Windows), ftp.mcom.com for Netscape (UNIX, Mac, Windows), or ftp.einet.net for EINET (Mac, Windows).

BOSTON CENTRAL UNIVERSITY LIBRARY

1. Start the WWW browser–executable program.
2. Open a new URL to http://www.ncbi.nlm.nih.gov/, then click to go to the WWW BLAST Search page.
3. Click in the pop-up menus to specify the search parameters.
4. Cut the query sequence file from the sequence data and paste it into the Sequence field.
5. Click the Perform Search button.

   The results will be displayed in the browser window.

EXAMPLE 1: USING BLASTN TO SEARCH A NUCLEOTIDE SEQUENCE DATABASE WITH A DNA SEQUENCE

In this example the cDNA sequence of human MLH1 (GenBank accession numbers U07343 and U07418) is used to search GenBank for related nucleotide sequences. MLH1 is a gene on chromosome 3p21-23 that is associated with hereditary nonpolyposis colon cancer (Bronner et al., 1994; Papadopoulos et al., 1994). The results of this search (BLASTN output) are shown in Figures 2.5.1 to 2.5.4.

Using BLAST Electronic Mail Server

```
-------------------------------------------------------------------
mail blast@ncbi.nlm.nih.gov
Subject:

PROGRAM blastn
DATALIB genbank
BEGIN
>HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds
GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG
TTATTCGGCGGCTGGACGAGACAGTGGTGAACCGCATCGCGGCGGGGGAAGTTATCCAGC
```

38 lines omitted for brevity.

```
ACCAACATAAGTGTTGGTAGCACTTAAGACTTATACTTGCCTTCTGATAGTATTCCTTTA
TACACAGTGGATTGATTATAAATAAATAGATGTGTCTTAACAT
-------------------------------------------------------------------
```

Using BLAST Command-Line Client Program

The command-line equivalent of the above BLASTN search is:

```
-------------------------------------------------------------------
my_computer> blastn genbank hmlh1.seq > hmlh1_genbank.blastn
-------------------------------------------------------------------
```

where my_computer is the command prompt on your system, hmlh1.seq is a file containing the sequence in FASTA format, and the > symbol instructs UNIX to save the output in a file named hmlh1_genbank.blastn

Discussion

The hit list produced by a BLASTN search (Fig. 2.5.2) is the most important part of the output: it is what researchers visually scan in their initial evaluation of the results. Each line of the hit list is composed of five fields (an additional field indicating the reading frame is included when BLASTX, TBLASTN, or TBLASTX is used). The first field contains a database designation, accession number, and locus name for the matched sequence separated by vertical bars. So, for the first hit, the database designation is gb (for

BOU IASRI CENTRAL UNIVERSITY LIBRARY

```
Trying cruncher... connected
National Center for Biotechnology Information (NCBI)

Experimental GENINFO(R) BLAST Network Service (Cruncher)

Mon Mar 27 12:27:33 EST 1995, Up 14 days, 10:04, load: 9.50, 8.01, 6.09

PEPTIDE SEQUENCE DATABASES
 nr  Non-redundant PDB+SwissProt+PIR+SPUpdate+GenPept+GPUpdate, updated daily
     for efficient, complete searches of the five component databases:
   pdb        Brookhaven Protein Data Bank, April 1994 Release
   swissprot  SWISS-PROT Release 31.0, March 1995
   pir        PIR Release 43.0 (complete), December 31, 1994
   spupdate   SWISS-PROT cumulative weekly update to the major release
   genpept    CDS translations from GenBank(R) Release 87.0, February 15, 1995
   gpupdate   cumulative daily updates to the major release of genpept
 kabatpro   Kabat Sequences of Proteins of Immunological Interest Release 5.0,
            August 1992
 tfd        TFD transcription factor (protein) database Release 7.0, June 1993
 alu *      Translations of select Alu repeats from REPBASE

NUCLEOTIDE SEQUENCE DATABASES
 nr    Non-redundant PDB+GBUpdate+GenBank+EmblUpdate+EMBL, updated daily
       for efficient, complete searches of the four component databases:
   pdb        Brookhaven Protein Data Bank, April 1994 Release
   genbank    GenBank(R) Release 87.0 (no daily updates), February 15, 1995
   gbupdate   GenBank(R) cumulative daily updates to the major release
   embl       EMBL Data Library, Release 41.0, December 1994
   emblu      EMBL Data Library cumulative daily updates to the major release
 vector     Vector subset of GenBank(R), February 3rd, 1995
 alu *+     Select Alu repeats from REPBASE
 kabatnuc   Kabat Sequences of Nucleic Acid of Immunological Interest
            Release 5.0, August 1992
 epd        Eukaryotic Promoter Database Release 35, June 1993
 dbest +    Database of Expressed Sequence Tags (cumulative daily update)
 dbsts +    Database of Sequence Tagged Sites Release 1.5, October 26, 1994

  * Databases that are not accessible through the NCBI Retrieve E-mail server.
  + The TBLASTX program is restricted to searching these databases.
================================================================================
 You can obtain the BLAST documentation files, send a message consisting of
 just the word ``help'' (without the quotes) to: blast@ncbi.nlm.nih.gov
 Last modification dates: January 25th 95 for the E-mail server help, January
 19th 94 for the BLAST manual and March 18th 95 for the BLAST FAQ.
================================================================================
 For a free subscription to "NCBI News", the NCBI newsletter, send a request
 along with your name and postal mailing address to: info@ncbi.nlm.nih.gov
================================================================================
 A new GenBank sequence submission tool, called BankIt, is now available
 through the NCBI's home page on the World Wide Web. The URL is
 http://www.ncbi.nlm.nih.gov/
================================================================================

BLASTN 1.4.7MP [19-Dec-94] [Build 16:54:55 Mar 16 1995]

Reference:  Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers,
and David J. Lipman (1990).  Basic local alignment search tool.  J. Mol. Biol.
215:403-10.

Notice:  this program and its default parameter settings are optimized to find
nearly identical sequences rapidly.  To identify weak similarities encoded in
nucleic acid, use BLASTX, TBLASTN or TBLASTX.
```

Figure 2.5.1 Contents of return message from BLAST network server containing BLASTN search results (see Example 1). It includes items to inform the user about databases available for searching and messages concerning other topics such as the status and availability of documentation.

BCU IASI CENTRAL UNIVERSITY LIBRARY

```
Query=  gi|466461|gb|U07418|HSHMLHI Human DNA mismatch re...
        (2503 letters)

Database:  GenBank(R) Release 87
           269,478 sequences; 248,499,214 total letters.
Searching.....................................................done

                                                                 Smallest
                                                                   Sum
                                                          High  Probability
Sequences producing High-scoring Segment Pairs:           Score  P(N)      N

gb|U07418|HSHMLHI     Human DNA mismatch repair (hmlh1) mR... 12515  0.0       1
gb|U07343|HSU07343    Human DNA mismatch repair protein ho... 12406  0.0       1
gb|T28624|T28624      EST49427 Homo sapiens cDNA 3' end si...  2032  5.3e-161  1
26 hits omitted for brevety
```

Figure 2.5.2 Information regarding the query sequence and database searched followed by the hit list for the BLASTN search in Figure 2.5.1.

GenBank), the accession number is U07418, and the locus name is HSHMLH1. The second field contains a brief textual description of the sequence, also known as the definition line. The content of this definition line varies between and within databases but usually includes information on the organism from which the sequence was derived, the type of sequence (e.g., mRNA or DNA), and some clue about function or phenotype. The definition line is often truncated in the hit list to keep the display compact but is fully displayed in association with the local alignments. The third field contains the alignment score, based on the scoring matrix used, for the best ungapped local alignment ("high-scoring segment pair" or HSP) between the query sequence and the database sequence. The fourth field contains a *p* value indicating the probability that the match could have occurred by chance and the fifth field indicates the number of HSPs in the set that resulted in the lowest *p* value. The third section of the output (Fig. 2.5.3) contains the sequence alignments prefaced by more complete descriptions of the database sequences and statistical results. Alignments found with the BLAST algorithm are ungapped. Several statistics describe each HSP: the raw alignment score; the raw score converted to bits of information; the number of times this match or a better one might occur merely by chance; the *p* value (probability in the range 0 to 1) of observing such a match; the number and fraction of total residues in the HSP that are identical; the number and fraction of residues for which the substitution scores have positive values; and, for nucleotide sequences, the strand of the query that resulted in the match. Between the two lines of Query and database (Sbjct) sequence is a line indicating by the symbol "|" the specific nucleotides that are identical; a blank space is left between those that are nonidentical. (Note that the convention is different for protein sequence alignments; see Sample 1 below). The section following the alignments at the end of the output (Fig. 2.5.4) contains a summary of the run parameters and statistics. Note that this search of 269,478 sequences using a 2503-nucleotide query took only 25 seconds (Total cpu time). The summary indicates that the BLAST results are complete.

Although it is not relevant to this example, BLASTN automatically checks the complementary strand of the query sequence. Note that BLASTN is optimized for speed, not sensitivity, and it is not intended for finding distantly related coding sequences (for comparison see Example 2). The small DNA alphabet (four letters) compared with proteins (twenty letters) and the degeneracy of the genetic code makes searches using the encoded protein sequence much more sensitive, as illustrated with the BLASTX and BLASTP searches below.

BU IACS LIBRARY UNIVERSITY LIBRARY

```
>gb|U07418|HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds.
           Length = 2503

self-alignment omitted for brevity

>gb|U07343|HSU07343 Human DNA mismatch repair protein homolog (hMLH1) mRNA,
           complete cds.
           Length = 2484

self-alignment omitted for brevity

>gb|T28624|T28624 EST49427 Homo sapiens cDNA 3' end similar to DNA mismatch
           repair protein hmlh1 (HT:3337).
           Length = 438

  Minus Strand HSPs:

 Score = 2032 (561.5 bits), Expect = 5.3e-161, P = 5.3e-161
 Identities = 408/411 (99%), Positives = 408/411 (99%), Strand = Minus / Plus

Query:  2503 ATGTTAAGACACATCTATTTATTTATAATCAATCCACTGTGTATAAAGGAATACTATCAG 2444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:     1 ATGTTAAGACACATCTATTTATTTATAATCAATCCACTGTGTATAAAGGAATACTATCAG 60

Query:  2443 AAGGCAAGTATAAGTCTTAAGTGCTACCAACACTTATGTTGGTACACTTTGTATATCACA 2384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:    61 AAGGCAAGTATAAGTCTTAAGTGCTACCAACACTTATGTTGGTACACTTTGTATATCACA 120

Query:  2383 CTTTGATACAACACTTTGTATCGGAATACAGAGAAAGAAGAACACATCCCACAGTGCATA 2324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:   121 CTTTGATACAACACTTTGTATCGGAATACAGAGAAAGAAGAACACATCCCACAGTGCATA 180

Query:  2323 AATAACCATATTTAACACCTCTCAAAGACTTTGTATAGATCAGGCAGGTTAGCAAGCTGC 2264
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:   181 AATAACCATATTTAACACCTCTCAAAGACTTTGTATAGATCAGGCAGGTTAGCAAGCTGC 240

Query:  2263 AGGATATTTCCATCTTCTGTGAAATGTTTAGGAGGCAGAATGTGTGAGCGCAAGGCTTTA 2204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:   241 AGGATATTTCCATCTTCTGTGAAATGTTTAGGAGGCAGAATGTGTGAGCGCAAGGCTTTA 300

Query:  2203 TAGACAATGTGTTCCACAGTCCACTTCCAGGAGTTTGGAATGGAGCCAGGCACTTCACTC 2144
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct:   301 TAGACAATGTGTTCCACAGTCCACTTCCNGGAGTTTGGAATGGAGCCAGGCACTTCACTC 360

Query:  2143 TGCTGGCCTGAGAGGGTCGACTCCTCAGATATGTACTGCTTCCGGATGGAA 2093
             |||||||||||||||||| |||||||||||||||||||||||||||| |
Sbjct:   361 TGCTGGCCTGAGAGGGTCGNCTCCTCAGATATGTACTGCTTCCGGATGGGA 411
```

Figure 2.5.3 Sequence alignments prefaced by more complete descriptions of the database sequences and statistical results for the BLASTN search in Figure 2.5.1. The two top (self) alignments from Figure 2.5.2 are omitted for brevity.

For proteins, it is also possible to take into account conservative amino acid substitutions (see Scoring Systems and Alignment Statistics in Commentary).

The BLAST programs have many options that alter the sensitivity, specificity, and speed of the search and specify different scoring systems, alternate genetic codes, etc. The examples show only the simplest cases. We illustrate the filter option (-filter command; see Example 3); a complete list of options and parameters is described in the BLAST help file. Note that there are certain restrictions to the list of available options provided by the BLAST e-mail server (see the e-mail help file). Options for the BLAST programs (Table 2.5.1) are preset to certain default parameters that are generally most useful. For tips on altering these parameters to address special situations, see Commentary.

BCU MEDICAL CENTRAL UNIVERSITY LIBRARY
CHINE

```
Parameters:

  V=500
  B=250
  -ctxfactor=2.00
  E=10

  Query                              -----  As Used  -----     -----  Computed  ----
  Strand MatID Matrix name          Lambda     K        H      Lambda     K        H
   +1       0   +5,-4               0.192   0.173    0.357    same     same     same
   -1       0   +5,-4               0.192   0.173    0.357    same     same     same

  Query
  Strand MatID  Length  Eff.Length   E      S  W   T  X     E2  S2
   +1       0    2503        2503    10.  125 11 N/A 73   0.024 82
   -1       0    2503        2503    10.  125 11 N/A 73   0.024 82


Statistics:
  Query            Expected            Observed           HSPs       HSPs
  Strand MatID  High Score          High Score         Reportable  Reported
   +1       0   132 (36.5 bits)   12515 (3458.1 bits)     27          27
   -1       0   132 (36.5 bits)    2037 (562.9 bits)      3           3

  Query          Neighborhd  Word       Excluded    Failed   Successful  Overlaps
  Strand MatID    Words      Hits         Hits     Extensions Extensions  Excluded
   +1       0      2496      201044       18102     180110      2902         1
   -1       0      2496      183944       15029     166528      2440         0

  Database:  GenBank(R) Release 87
    Release date:  February 15, 1995
    Posted date:  5:37 PM EST Feb 9, 1995
  # of letters in database:  248,499,214
  # of sequences in database:  269,478
  # of database sequences satisfying E:  29
  No. of states in DFA:  255 (255 KB)
  Total size of DFA:  313 KB (320 KB)
  Time to generate neighborhood:  0.04u 0.02s 0.06t  Real: 00:00:00
  No. of processors used:  3
  Time to search database:  37.94u 4.08s 42.02t  Real: 00:00:24
  Total cpu time:  38.09u 4.18s 42.27t  Real: 00:00:25
```

Figure 2.5.4 Summary of the run parameters and statistics for the BLASTN search in Figure 2.5.1.

EXAMPLE 2: USING BLASTX TO SEARCH A PROTEIN SEQUENCE DATABASE WITH OPEN READING FRAMES DERIVED FROM CONCEPTUAL TRANSLATIONS OF A DNA QUERY SEQUENCE

Suppose that for the DNA sequence in Example 1 the precise location of the translation initiation or termination codons was not known. Or suppose that the sequence was from genomic DNA and it was not clear whether it contained any coding sequences at all. Perhaps the data were derived from single-pass automated sequencing and contained some frameshift errors or nonsense mutations causing premature termination. In compact genomes such as those of yeast and prokaryotes, where the intergenic distances are small, there may be another gene upstream or downstream of the gene of interest. These are all "real world" contingencies that BLASTX was designed to handle (Gish and States, 1993).

In this example, the database designation nr refers to NCBI's non-redundant sequence database. The output of this search is shown in Sample 1 and in Figure 2.5.5.

BRUNEL UNIVERSITY LIBRARY

Using BLAST Electronic Mail Server

```
-----------------------------------------------------------------
mail blast@ncbi.nlm.nih.gov
Subject:

PROGRAM blastx
DATALIB nr
BEGIN
>HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds
GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG
TTATTCGGCGGCTGGACGAGACAGTGGTGAACCGCATCGCGGCGGGGGAAGTTATCCAGC
```

38 lines omitted for brevity.

```
ACCAACATAAGTGTTGGTAGCACTTAAGACTTATACTTGCCTTCTGATAGTATTCCTTTA
TACACAGTGGATTGATTATAAATAAATAGATGTGTCTTAACAT
-----------------------------------------------------------------
```

Using BLAST Command-Line Client Program

The command-line equivalent of the above BLASTX search is:

```
-----------------------------------------------------------------
my_computer> blastx nr hmlh1.seq > hmlh1_nr.blastx &
-----------------------------------------------------------------
```

In this command-line example, the ampersand (&) instructs UNIX to run this job as a background process. Because BLASTX is translating the nucleotide query in all six reading frames, this search will generally take longer than a BLASTN search, depending on the size of the database, and it is wise to run such searches in the background. This caution also applies to the TBLASTN program (Example 5).

Sample 1: Output from Example 2 Search

```
>gp|U07418|HSHMLHI_1 hmlh1 gene product [Homo sapiens]
           Length = 756
```

Self-alignment omitted for brevity.

```
>sp|P40692|MLH1_HUMAN MUTL PROTEIN HOMOLOG 1 (DNA MISMATCH REPAIR PROTEIN
           MLH1). >pir|S43085|S43085 MLH1 protein - human
           >gp|U07343|HSU07343_1 hMLH1 [Homo sapiens]
           Length = 756
```

Self-alignment omitted for brevity.

```
>gp|U17857|HSMLH1S19_1 hMLH1 gene product [Homo sapiens]
           Length = 752
```

Self-alignment omitted for brevity.

```
>sp|P38920|MLH1_YEAST MUTL PROTEIN HOMOLOG 1 (DNA MISMATCH REPAIR PROTEIN
           MLH1). >pir|S48890|S48890 MLH1 protein - yeast (Saccharomyces
           cerevisiae) >gp|U07187|SCMLHI_1 Mlh1p [Saccharomyces cerevisiae]
           Length = 769

Plus Strand HSPs:

Score = 676 (311.0 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 128/222 (57%), Positives = 170/222 (76%), Frame = +3
```

BOSTON/CENTRAL UNIVERSITY LIBRARY

```
Query:    63 IRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTG 242
             I+ LD +VVN+IAAGE+I  P NA+KEM+EN +DA +T I ++VKEGG+K++QI DNG+G
Sbjct:     5 IKALDASVVNKIAAGEIIISPVNALKEMMENSIDANATMIDILVKEGGIKVLQITDNGSG 64

Query:   243 IRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYR 422
             I K DL I+CERFTTSKLQ FEDL+ I TYGFRGEALASISHVA VT+TTK  + +CA+R
Sbjct:    65 INKADLPILCERFTTSKLQKFEDLSQIQTYGFRGEALASISHVARVTVTTKVKEDRCAWR 124

Query:   423 ASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHN 602
              SY++GK+   PKP AG  GT I VEDLF+NI +R +AL++ ++EY KIL+VVGRY++H+
Sbjct:   125 VSYAEGKMLESPKPVAGKDGTTILVEDLFFNIPSRLRALRSHNDEYSKILDVVGRYAIHS 184

Query:   603 AGISFSVKKQGETVADVRTLPNASTVDNIRSVFGNAVSRELI 728
               I FS KK G++   +   P+ +  D IR+VF  +V+  LI
Sbjct:   185 KDIGFSCKKFGDSNYSLSVKPSYTVQDRIRTVFNKSVASNLI 226

 Score = 217 (99.8 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 41/92 (44%), Positives = 60/92 (65%), Frame = +3

Query:   795 SVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTK 974
             S K    + FIN+RLV    LR+A+ +VY+ YLPK   PF+YL + I P  VDVNVHPTK
Sbjct:   252 SKKSISLIFFINNRLVTCDLLRRALNSVYSNYLPKGFRPFIYLGIVIDPAAVDVNVHPTK 311

Query:   975 HEVHFLHEESILERVQQHIESKLLGSNSSRMY 1070
              EV FL ++ I+E++   + ++L   ++SR +
Sbjct:   312 REVRFLSQDEIIEKIANQLHAELSAIDTSRTF 343

 Score = 131 (60.3 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 23/49 (46%), Positives = 33/49 (67%), Frame = +3

Query:  1947 LIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESLSKECAMFY 2093
             L  LPLL+  Y+P L  LP FI RL  EV+W++E+EC + + +E A+ Y
Sbjct:   649 LKSLPLLLKGYIPSLVKLPFFIYRLGKEVDWEDEQECLDGILREIALLY 697

 Score = 93 (42.8 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 21/52 (40%), Positives = 29/52 (55%), Frame = +3

Query:  1656 ALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLAL 1811
             A  QH  KL+L++   +  ELFYQI + DFANFG + L       D+ +  L
Sbjct:   549 AAIQHDLKLFLIDYGSVCYELFYQIGLTDFANFGKINLQSTNVSDDIVLYNL 600

 Score = 81 (37.3 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 13/18 (72%), Positives = 18/18 (100%), Frame = +3

Query:  2256 NILQLANLPDLYKVFERC 2309
             +++++ANLPDLYKVFERC
Sbjct:   752 DVVEIANLPDLYKVFERC 769

 Score = 76 (35.0 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 17/49 (34%), Positives = 30/49 (61%), Frame = +3

Query:  1542 INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYL 1688
             +NLTS+  L+E+++  H  L ++  N ++VG V+ +  LA  Q  L L
Sbjct:   509 VNLTSIKKLREKVDDSIHRELTDIFANLNYVGVVDEERRLAAIQHDLKL 557

 Score = 49 (22.5 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 6/22 (27%), Positives = 14/22 (63%), Frame = +3

Query:  2187 VEHIVYKALRSHILPPKHFTED 2252
             +EH+++  ++   L P+H  +D
Sbjct:   731 LEHVLFPCIKRRFLAPRHILKD 752

 Score = 42 (19.3 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 8/26 (30%), Positives = 16/26 (61%), Frame = +3
```

RIJ IASI / CENTRAL UNIVERSITY LIBRARY

```
Query:  1866 EYIVEFLKKKAEMLADYFSLEIDEEG 1943
             E I+  +   + ML +Y+S+E+  +G
Sbjct:   614 EKIISKIWDMSSMLNEYYSIELVNDG 639

 Score = 41 (18.9 bits), Expect = 9.2e-195, Sum P(9) = 9.2e-195
 Identities = 8/20 (40%), Positives = 13/20 (65%), Frame = +3

Query:   765 MNGYISNANYSVKKCIFLLF 824
             ++G + N N+  KK I L+F
Sbjct:   241 VDGKVCNLNFISKKSISLIF 260

 Score = 39 (17.9 bits), Expect = 2.5e-34, Sum P(7) = 2.5e-34
 Identities = 9/27 (33%), Positives = 14/27 (51%), Frame = +3

Query:  1272 IVTEDKTDISSGRARQQDEEMLELPAP 1352
             + T+ K D  + R    + +MLE P P
Sbjct:   112 VTTKVKEDRCAWRVSYAEGKMLESPKP 138
```

Remaining alignments omitted for brevity.

Discussion

There are several differences between the hit list from a BLASTX search (Fig. 2.5.5) and the list from a BLASTN search (Fig. 2.5.2). First, because NCBI's nr nonredundant database was searched, the database designations are variable, reflecting the source databases used to build nr. For example, gp indicates a GenPept translation of a GenBank sequence and sp indicates Swiss-Prot was the source database. More complete descriptions of the database cross-references are included in the long header lines associated with the alignments in Sample 1. The second difference is the presence of the reading frame

```
Query=  gi|466461|gb|U07418|HSHMLHI Human DNA mismatch re...
        (2503 letters)

  Translating both strands of query sequence in all 6 reading frames

Database:  Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate
           143,811 sequences; 41,308,423 total letters.
Searching.....................................................done
```

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Sum Probability P(N)	N
gp|U07418|HSHMLHI_1	hmlh1 gene product [Homo sapiens]	+3	3869	0.0	1
sp|P40692|MLH1_HUMAN	MUTL PROTEIN HOMOLOG 1 (DNA MISM...	+3	3868	0.0	1
gp|U17857|HSMLH1S19_1	hMLH1 gene product [Homo sapiens]	+3	3595	0.0	2
sp|P38920|MLH1_YEAST	MUTL PROTEIN HOMOLOG 1 (DNA MISM...	+3	676	9.2e-195	9
sp|P23367|MUTL_ECOLI	DNA MISMATCH REPAIR PROTEIN MUTL...	+3	409	1.9e-57	2
sp|P14161|MUTL_SALTY	DNA MISMATCH REPAIR PROTEIN MUTL...	+3	399	7.9e-57	4
sp|P14160|HEXB_STRPN	DNA MISMATCH REPAIR PROTEIN HEXB...	+3	261	1.2e-47	5
gp|U13696|HSU13696_1	hPMS2 gene product [Homo sapiens]	+3	278	2.1e-38	6
gp|U14658|HSU14658_1	hPMS2 gene product [Homo sapiens]	+3	278	2.1e-38	6
sp|P14242|PMS1_YEAST	DNA MISMATCH REPAIR PROTEIN PMS1...	+3	230	5.2e-35	7
pir|A33590|A33590	PMS1 protein - yeast (Saccharomy...	+3	230	2.1e-33	6
gp|U13695|HSU13695_1	hPMS1 gene product [Homo sapiens]	+3	268	8.3e-33	3
sp|P32583|SR40_YEAST	SUPRESSOR PROTEIN SRP40. >pir|S3...	+3	72	0.023	2
gp|L11275|YSCSRP40X_1	SRP40 gene product [Saccharomyce...	+3	72	0.023	2
gp|L26379|BPL26379N_1	protein 17 [Bacteriophage T7] >g...	+3	45	0.20	2
sp|P08640|AMYH_YEAST	GLUCOAMYLASE S1/S2 PRECURSOR (EC...	+3	69	0.31	4
pir|A49902|A49902	protein-histidine kinase (EC 2.7...	+3	80	0.32	1
sp|P27476|NSR1_YEAST	NUCLEAR LOCALIZATION SEQUENCE BI...	+3	78	0.50	1
pir|A44937|A44937	kinetoplast-associated protein -...	+3	78	0.54	1
gp|L26367|BPL26367J_1	protein 17 [Bacteriophage T7] >g...	+3	45	0.54	2

BLAST shows the top 100 matches by default; only the first 20 are included to save space.

Figure 2.5.5 Contents of return message from BLAST e-mail server containing BLASTX search results from Example 2. Hit list. Text in italic is added by the author.

UCSF CENTRAL UNIVERSITY LIBRARY

in the third field of this hit list. +3 indicates the third reading frame on the query sequence and −1 designates the first reading frame on the complementary strand. The main difference in the alignments (Sample 1) is the labeling of exact matches and conservative amino acid substitutions. An amino acid identity between two aligned sequences is indicated by simply repeating the single-letter code symbol for that residue. Conservative substitutions are indicated by a plus (+) sign; an amino acid substitution is defined as conservative if it has a replacement score of ≥0 in the matrix used for the alignment. In Example 1, there happened to be only a single HSP for each query/database match. In the present example most matches consist of multiple HSPs, and these are reported not in order of their occurrence in the linear sequences of the proteins, but rather in order of their degree of similarity. This is one of the difficulties for beginners (who are used to seeing more extensive alignments with gaps) in trying to understand and interpret the BLAST output. The best of both worlds is provided by the NetBLAST graphical interface (see Example 6), which not only displays the global relationships schematically (Fig. 2.5.10B) but also provides point-and-click access to the textual representations of all reported local alignments. Because all six reading frames of the query had to be tested against the database, this search took several times longer than the BLASTN search in Example 1.

BLASTX is very useful for analyzing preliminary sequence data containing potential errors (Gish and States, 1993) and is essential for analyzing expressed sequence tags (ESTs; Boguski et al., 1993, 1994) or genomic sampling sequence data (Smith et al., 1994). Because these data are usually generated by single-pass sequencing, the error rate is ~1% to 3%. Moreover, the coding potential is not known in advance, making six-frame translations necessary.

Note that, in this search involving conceptual translations of the nucleotide query, the "standard" or "universal" genetic code is used by default. For scoring the amino acid alignments, the BLOSUM62 matrix is the default system. As noted, either of these parameters can be changed if appropriate. For further information see Commentary.

EXAMPLE 3: USING BLASTP TO SEARCH A PROTEIN SEQUENCE DATABASE WITH A PROTEIN SEQUENCE QUERY USING THE FILTER OPTION TO MASK LOW-COMPLEXITY SUBSEQUENCES

In this case we will again search the nr database, this time using the translated protein sequence of MLH1 (Example 1) as the query and "masking" low-complexity subsequence in the query using the seg algorithm (see below). Note that the manner or syntax for specifying this filter option is slightly different in the e-mail and command-line versions. The results of this search, showing the effects of query masking, are given in Sample 2 and in Figure 2.5.6.

Using BLAST Electronic Mail Server

```
------------------------------------------------------------------
mail blast@ncbi.nlm.nih.gov
Subject:

PROGRAM blastp
DATALIB nr
FILTER seg
BEGIN
>HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds.
MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKS
TSIQVIVKEGGLKLIQIQDNGT
```

Lines omitted for brevity.

RJ IAS CENTRAL UNIVERSITY LIBRARY

```
FYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNI
LQLANLPDLYKVFERC
------------------------------------------------------------
```

Using BLAST Command-Line Client Program

The command-line equivalent of the above BLASTX search is:

```
------------------------------------------------------------
my_computer> blastp nr hmlh1.aa -filter seg > hmlh1_nr.blastp
------------------------------------------------------------
```

Sample 2: Output from Example 3 Search

```
>gp|U07418|HSHMLHI_1 hmlh1 gene product [Homo sapiens]
           Length = 756
 Score = 3735 (1703.0 bits), Expect = 0.0, P = 0.0
 Identities = 730/756 (96%), Positives = 730/756 (96%)

Query:     1 MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ 60
             MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ
Sbjct:     1 MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ 60
```

Portion of self-alignment omitted for brevity. Note the serine- and threonine-rich low-complexity subsequence replaced by "X."

```
Query:   361 KXXXXXXXXXXXXXXXDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS 420
             K                DKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS
Sbjct:   361 KSTTSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS 420

Query:   421 SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHRXXXXXX 480
             SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHR
Sbjct:   421 SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVE 480

Query:   481 XXXXXXRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL 540
                  RKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL
Sbjct:   481 MVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL 540
```

Remainder of self-alignment omitted for brevity.

```
>sp|P40692|MLH1_HUMAN MUTL PROTEIN HOMOLOG 1 (DNA MISMATCH REPAIR PROTEIN
           MLH1). >pir|S43085|S43085 MLH1 protein - human
           >gp|U07343|HSU07343_1 hMLH1 [Homo sapiens]
           Length = 756

 Score = 3734 (1702.5 bits), Expect = 0.0, P = 0.0
 Identities = 729/756 (96%), Positives = 730/756 (96%)
```

Alignment omitted for brevity.

The following alignment is identical to that produced by the BLASTX program in Sample 1. However, note the variation in numbering. Here amino acid residues 8 through 229 correspond to nucleotides 39 through 623 in Sample 1. Sample 1 also includes the translation frame information, e.g., "Frame = +3."

```
>gp|U17857|HSMLH1S19_1 hMLH1 gene product [Homo sapiens]
           Length = 752

 Score = 3461 (1578.1 bits), Expect = 0.0, Sum P(2) = 0.0
 Identities = 680/707 (96%), Positives = 681/707 (96%)

Query:     1 MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ 60
             MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ
Sbjct:     1 MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQ 60
```

BCU INSTITUTE CENTRAL UNIVERSITY LIBRARY

```
Query:    61 IQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTA 120
             IQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTA
Sbjct:    61 IQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTA 120

Query:   121 DGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEYGKILEVV 180
             DGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEYGKILEVV
Sbjct:   121 DGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEYGKILEVV 180

Query:   181 GRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSVFGNAVSRELIEIGCEDKTLAF 240
             GRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRS+FGNAVSRELIEIGCEDKTLAF
Sbjct:   181 GRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFGNAVSRELIEIGCEDKTLAF 240

Query:   241 KMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISP 300
             KMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISP
Sbjct:   241 KMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISP 300

Query:   301 QNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMV 360
             QNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMV
Sbjct:   301 QNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMV 360

Query:   361 KXXXXXXXXXXXXXXDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS 420
             K              DKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS
Sbjct:   361 KSTTSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDIS 420

Query:   421 SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHRXXXXXX 480
             SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHR
Sbjct:   421 SGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVE 480

Query:   481 XXXXXXRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL 540
                   RKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL
Sbjct:   481 MVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWAL 540

Query:   541 AQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEE 600
             AQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEE
Sbjct:   541 AQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEE 600

Query:   601 DGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL 660
             DGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL
Sbjct:   601 DGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL 660

Query:   661 ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGS 707
             ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGS
Sbjct:   661 ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGS 707

 Score = 266 (121.3 bits), Expect = 0.0, Sum P(2) = 0.0
 Identities = 47/49 (95%), Positives = 48/49 (97%)

Query:   708 IPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC 756
             +P SWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC
Sbjct:   704 VPGSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC 752

>sp|P38920|MLH1_YEAST MUTL PROTEIN HOMOLOG 1 (DNA MISMATCH REPAIR PROTEIN
          MLH1). >pir|S48890|S48890 MLH1 protein - yeast (Saccharomyces
          cerevisiae) >gp|U07187|SCMLHI_1 Mlh1p [Saccharomyces cerevisiae]
          Length = 769

 Score = 676 (308.2 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 128/222 (57%), Positives = 170/222 (76%)

Query:     8 IRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTG 67
             I+ LD +VVN+IAAGE+I  P NA+KEM+EN +DA +T I ++VKEGG+K++QI DNG+G
Sbjct:     5 IKALDASVVNKIAAGEIIISPVNALKEMMENSIDANATMIDILVKEGGIKVLQITDNGSG 64

Query:    68 IRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYR 127
             I K DL I+CERFTTSKLQ FEDL+ I TYGFRGEALASISHVA VT+TTK  + +CA+R
Sbjct:    65 INKADLPILCERFTTSKLQKFEDLSQIQTYGFRGEALASISHVARVTVTTKVKEDRCAWR 124
```


BU IASI / CENTRALA UNIVERSITARA

```
Query:   128 ASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHN 187
             SY++GK+   PKP AG  GT I VEDLF+NI +R +AL++ ++EY KIL+VVGRY++H+
Sbjct:   125 VSYAEGKMLESPKPVAGKDGTTILVEDLFFNIPSRLRALRSHNDEYSKILDVVGRYAIHS 184

Query:   188 AGISFSVKKQGETVADVRTLPNASTVDNIRSVFGNAVSRELI 229
              I FS KK G++   +   P+ +  D IR+VF  +V+  LI
Sbjct:   185 KDIGFSCKKFGDSNYSLSVKPSYTVQDRIRTVFNKSVASNLI 226

 Score = 217 (98.9 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 41/92 (44%), Positives = 60/92 (65%)

Query:   252 SVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTK 311
             S K    + FIN+RLV    LR+A+ +VY+ YLPK   PF+YL + I P  VDVNVHPTK
Sbjct:   252 SKKSISLIFFINNRLVTCDLLRRALNSVYSNYLPKGFRPFIYLGIVIDPAAVDVNVHPTK 311

Query:   312 HEVHFLHEESILERVQQHIESKLLGSNSSRMY 343
              EV FL ++ I+E++   + ++L   ++SR +
Sbjct:   312 REVRFLSQDEIIEKIANQLHAELSAIDTSRTF 343

 Score = 131 (59.7 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 23/49 (46%), Positives = 33/49 (67%)

Query:   636 LIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESLSKECAMFY 684
             L  LPLL+  Y+P L  LP FI RL  EV+W++E+EC + + +E A+ Y
Sbjct:   649 LKSLPLLLKGYIPSLVKLPFFIYRLGKEVDWEDEQECLDGILREIALLY 697

 Score = 93 (42.4 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 21/52 (40%), Positives = 29/52 (55%)

Query:   539 ALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLAL 590
             A  QH  KL+L++   + ELFYQI + DFANFG + L        D+ +  L
Sbjct:   549 AAIQHDLKLFLIDYGSVCYELFYQIGLTDFANFGKINLQSTNVSDDIVLYNL 600

 Score = 81 (36.9 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 13/18 (72%), Positives = 18/18 (100%)

Query:   739 NILQLANLPDLYKVFERC 756
             +++++ANLPDLYKVFERC
Sbjct:   752 DVVEIANLPDLYKVFERC 769

 Score = 76 (34.7 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 17/49 (34%), Positives = 30/49 (61%)

Query:   501 INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYL 549
             +NLTS+  L+E++++  H  L ++  N ++VG V+ +  LA  Q  L L
Sbjct:   509 VNLTSIKKLREKVDDSIHRELTDIFANLNYVGVVDEERRLAAIQHDLKL 557

 Score = 49 (22.3 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 6/22 (27%), Positives = 14/22 (63%)

Query:   716 VEHIVYKALRSHILPPKHFTED 737
             +EH+++  ++   L P+H  +D
Sbjct:   731 LEHVLFPCIKRRFLAPRHILKD 752

 Score = 42 (19.2 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 8/26 (30%), Positives = 16/26 (61%)

Query:   609 EYIVEFLKKKAEMLADYFSLEIDEEG 634
             E I+  +   + ML +Y+S+E+  +G
Sbjct:   614 EKIISKIWDMSSMLNEYYSIELVNDG 639

 Score = 41 (18.7 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 8/20 (40%), Positives = 13/20 (65%)

Query:   242 MNGYISNANYSVKKCIFLLF 261
             ++G + N N+  KK I L+F
Sbjct:   241 VDGKVCNLNFISKKSISLIF 260
```

BCU IUPUI/CENTRAL UNIVERSITY LIBRARY

```
 Score = 39 (17.8 bits), Expect = 1.9e-35, Sum P(7) = 1.9e-35
 Identities = 9/27 (33%), Positives = 14/27 (51%)

Query:   411 IVTEDKTDISSGRARQQDEEMLELPAP 437
             + T+ K D  + R    + +MLE P P
Sbjct:   112 VTTKVKEDRCAWRVSYAEGKMLESPKP 138

 Score = 37 (16.9 bits), Expect = 1.2e-188, Sum P(9) = 1.2e-188
 Identities = 7/22 (31%), Positives = 13/22 (59%)

Query:   503 LTSVLSLQEEINEQGHEVLREM 524
             +TS LS  ++ N +G    R++
Sbjct:   409 ITSFLSSSQQFNFEGSSTKRQL 430

 Score = 36 (16.4 bits), Expect = 1.4e-200, Sum P(10) = 1.4e-200
 Identities = 6/17 (35%), Positives = 13/17 (76%)

Query:   381 HQMVRTDSREQKLDAFL 397
             +++VR D+ + K+ +FL
Sbjct:   397 NKLVRIDASQAKITSFL 413

 Score = 36 (16.4 bits), Expect = 6.3e-66, Sum P(9) = 6.3e-66
 Identities = 9/40 (22%), Positives = 20/40 (50%)

Query:    14 TVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKE 53
             TV N+  A  +I    + ++++    +D K  ++  I K+
Sbjct:   215 TVFNKSVASNLITFHISKVEDLNLESVDGKVCNLNFISKK 254

 Score = 35 (16.0 bits), Expect = 2.0e-140, Sum P(7) = 2.0e-140
 Identities = 6/23 (26%), Positives = 14/23 (60%)

Query:   209 NASTVDNIRSVFGNAVSRELIEI 231
             N +++  +R    +++ REL +I
Sbjct:   510 NLTSIKKLREKVDDSIHRELTDI 532
```

Discussion

There are two striking differences between the hit list produced by BLASTP (Fig. 2.5.6) and that of the BLASTX search with the same query (Fig. 2.5.5); both are due to masking of low-complexity subsequences using the SEG filter option. This hit list contains only 47 matches (compared with 356 matches), and also the identities of the matches are completely different (and inconsistent) following the ninth match in both lists. Low-complexity subsequences identified by the SEG algorithm are masked by replacing the amino acid codes with "X"; two such subsequences (residues 365 to 375 and 475 to 486) can be seen in the first alignment. BLAST ignores these X-masked segments. The first segment is proline-rich and the second segment has a high proportion of acidic—Glu (E) and Asp (D)—residues. These low-complexity subsequences represent less than 3% of the query sequence, yet their removal has a dramatic impact, shrinking the output by reducing the number of spurious matches and greatly improving the signal-to-noise ratio.

It cannot be stressed enough how useful it is to filter and mask low-complexity subsequences using the SEG and/or XNU algorithms and a thorough discussion of this important topic is included (see Low-Complexity Sequences and Masking Low-Complexity Sequences in Commentary). Masking procedures are absolutely essential when faced with analyzing the output of high-throughput sequencing where, for obvious practical reasons, human intervention in the sequence analysis process must be streamlined as much as possible.

RAJIV GANDHI UNIVERSITY LIBRARY

```
Query=  gi|466462|gp|U07418|HSHMLHI_1 hmlh1 gene product ...
        (756 letters)

Database:  Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate
           143,811 sequences; 41,308,423 total letters.
Searching.....................................................done

                                                                  Smallest
                                                                    Sum
                                                          High   Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)       N

gp|U07418|HSHMLHI_1   hmlh1 gene product [Homo sapiens]    3735  0.0          1
sp|P40692|MLH1_HUMAN  MUTL PROTEIN HOMOLOG 1 (DNA MISMATC...  3734  0.0          1
gp|U17857|HSMLH1S19_1 hMLH1 gene product [Homo sapiens]    3461  0.0          2
sp|P38920|MLH1_YEAST  MUTL PROTEIN HOMOLOG 1 (DNA MISMATC...   676  1.4e-200 10
sp|P23367|MUTL_ECOLI  DNA MISMATCH REPAIR PROTEIN MUTL. >...   409  6.4e-58   2
sp|P14161|MUTL_SALTY  DNA MISMATCH REPAIR PROTEIN MUTL. >...   399  1.3e-57   4
sp|P14160|HEXB_STRPN  DNA MISMATCH REPAIR PROTEIN HEXB. >...   261  1.5e-48   5
gp|U13696|HSU13696_1  hPMS2 gene product [Homo sapiens]     278  4.6e-40   7
gp|U14658|HSU14658_1  hPMS2 gene product [Homo sapiens]     278  4.6e-40   7
gp|U13695|HSU13695_1  hPMS1 gene product [Homo sapiens]     268  1.9e-33   3
sp|P14242|PMS1_YEAST  DNA MISMATCH REPAIR PROTEIN PMS1. >...   251  8.9e-31   4
pir|A33590|A33590     PMS1 protein - yeast (Saccharomyces...   251  8.9e-31   4
```

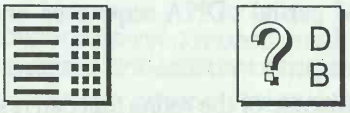
Figure 2.5.6 Contents of return message from BLAST e-mail server containing BLASTP search results from Example 3. Hit list.

EXAMPLE 4: USING THE WWW INTERFACE

Here we illustrate the use of the WWW BLAST interface by repeating the search described in Example 1 but with the nucleotide nonredundant database.

Figure 2.5.7 shows the WWW BLAST interface Search page. Pop-up menus (marked by the arrows) allow the user to define the kind of search needed by choosing the type of program, the database, and the BLAST options available through this interface. This example (Fig. 2.5.7) uses the default parameters. Click on each of the underlined words (e.g., Program, Expect, Cutoff) to see on-line documentation about the items. Insert comments about the sequence into the Comment field. Insert the sequence into the Query Sequence field simply by copying and pasting it from the sequence file. Then click on Perform Search to run the search. Once the search is done, the results are returned in the WWW browser window. Figure 2.5.8 shows the Output page with the hit list and the beginning of the first alignment.

Two features allow you to move back and forth in the BLAST output and are very useful for large BLAST output files. First, at the level of the hit list, clicking on the underlined part of one of the hit list lines moves you directly to the corresponding alignment. Second, each alignment has two buttons (see Fig. 2.5.8 and Sketch 2.5A): clicking on the left-hand button moves you back to the hit list; clicking on the right-hand button retrieves the entry of the matched sequence from the database and displays the entry on screen (see Fig. 2.5.9). This example shows the GenBank entry U07418 corresponding to a human mismatch repair mRNA. To move back to the BLAST Output page, click on the Back button of the WWW browser.



Sketch 2.5A

BCU IAS / CITY AS UNIVERSITY LIBRARY

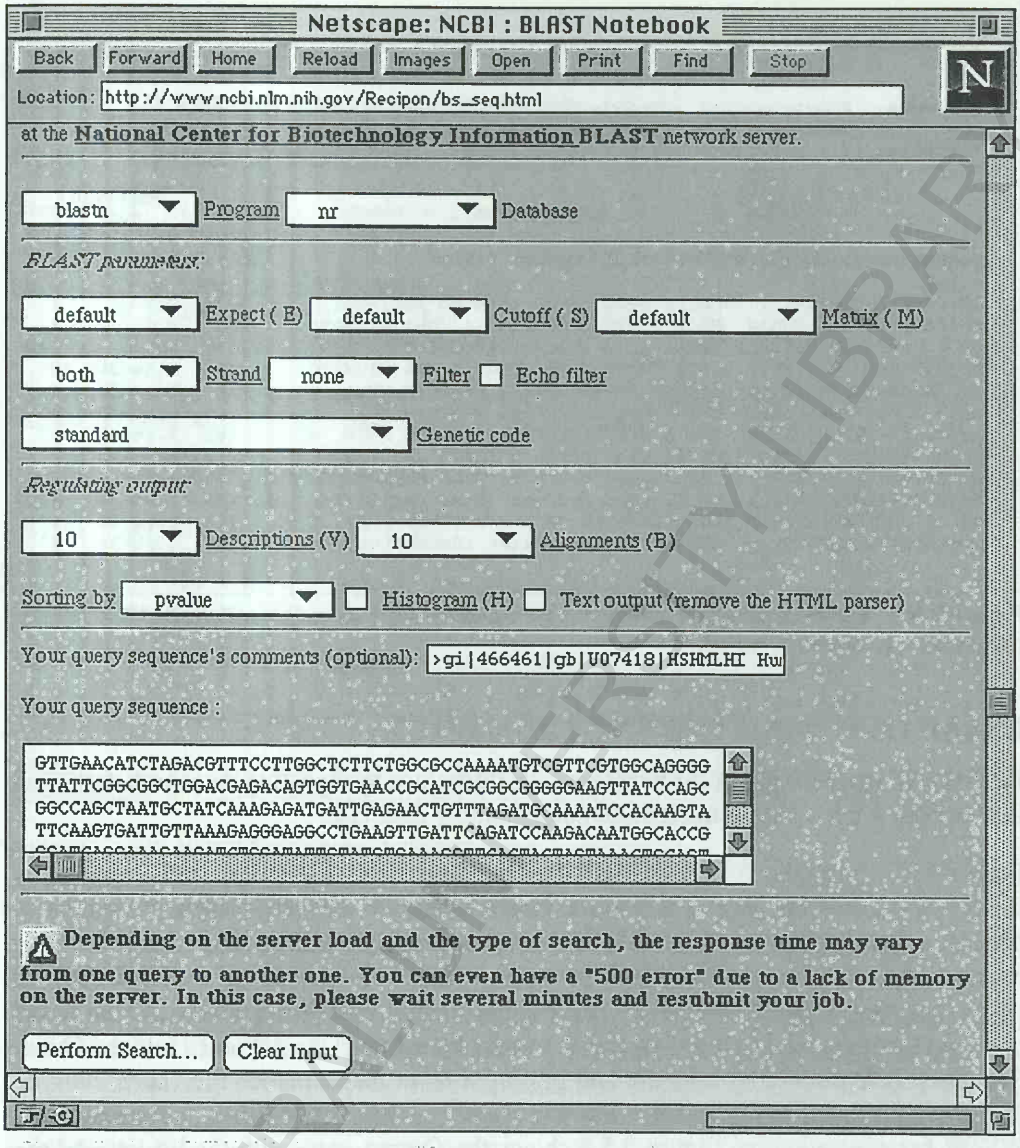


Figure 2.5.7 The WWW interface of the BLAST network server (see Example 4): Search page.

EXAMPLE 5: USING TBLASTN TO SEARCH A TRANSLATED NUCLEOTIDE DATABASE WITH A PROTEIN SEQUENCE

TBLASTN performs a function complementary to BLASTX (Example 2); i.e., whereas BLASTX translates the nucleotide query sequence in all six frames, then searches a protein database, TBLASTN takes a protein query sequence and compares it against six-frame translations of a nucleotide sequence database. As this latter procedure requires a great deal of computation, it should not be undertaken lightly on large nucleotide sequence databases. Nevertheless, the reality is that nucleotide sequence databases do contain errors and undocumented coding sequences making TBLASTN an important part of the database searching toolkit. TBLASTN is essential for searching protein queries against dbEST, a collection of partial cDNA sequences or "expressed sequence tags" (Boguski et al., 1993, 1994).

Choroideremia is an X-linked disease of the retina that can result in blindness (McKusick, 1994). The gene responsible encodes a 95-kDa protein that we will use as a query sequence.

BU IIAS CENTRAL UNIVERSITY LIBRARY

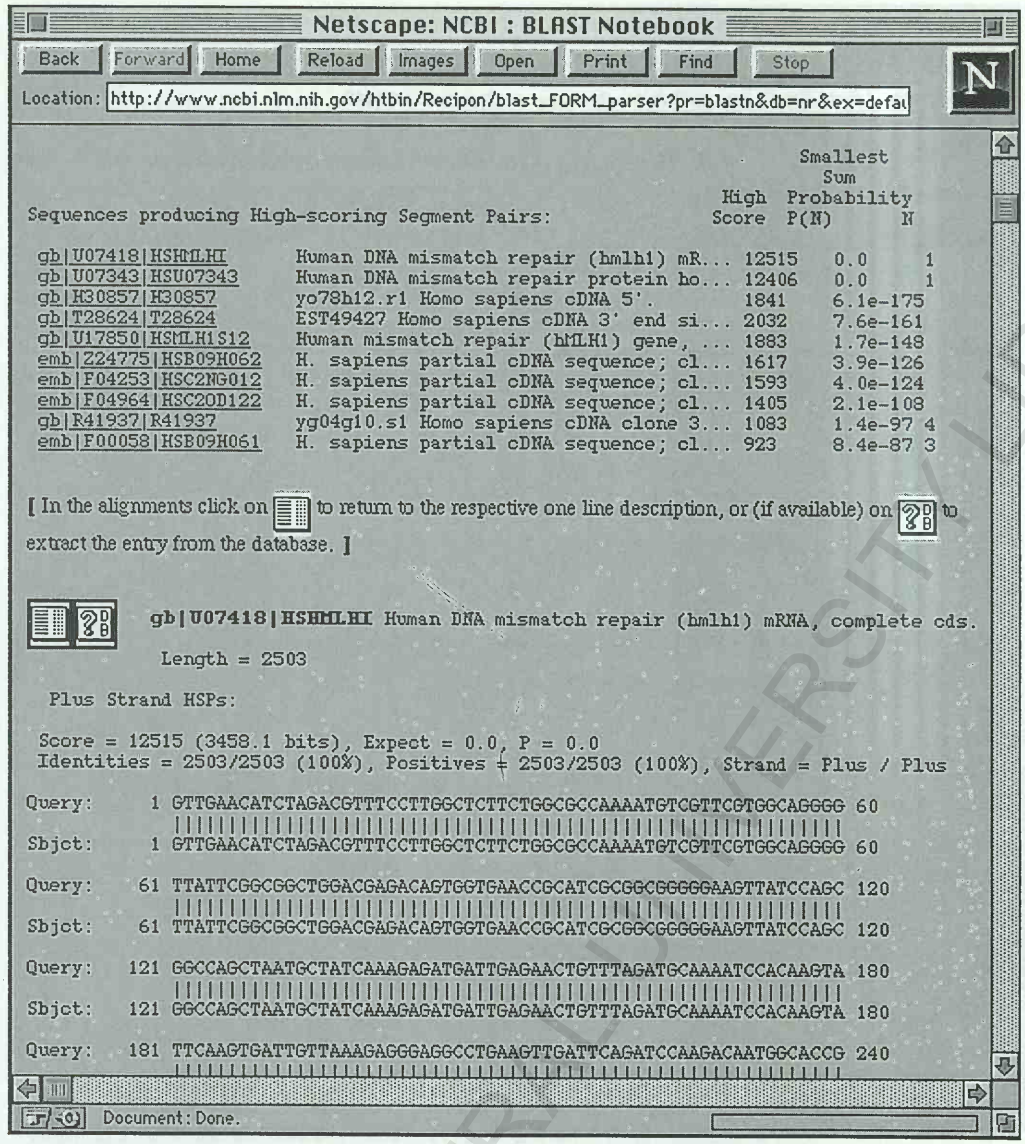

Netscape: NCBI : BLAST Notebook
Back
Forward
Home
Reload
Images
Open
Print
Find
Stop
Location: http://www.ncbi.nlm.nih.gov/htbin/Recipon/blast_FORM_parser?pr=blastn&db=nr&ex=defau
Sequences producing High-scoring Segment Pairs:
High Score
Smallest Sum Probability P(N)
N
gb|U07418|HSHMLHI Human DNA mismatch repair (hmlh1) mR... 12515 0.0 1
gb|U07343|HSU07343 Human DNA mismatch repair protein ho... 12406 0.0 1
gb|H30857|H30857 yo78h12.r1 Homo sapiens cDNA 5'. 1841 6.1e-175
gb|T28624|T28624 EST49427 Homo sapiens cDNA 3' end si... 2032 7.6e-161
gb|U17850|HSMLH1S12 Human mismatch repair (hMLH1) gene, ... 1883 1.7e-148
emb|Z24775|HSB09H062 H. sapiens partial cDNA sequence; cl... 1617 3.9e-126
emb|F04253|HSC2NG012 H. sapiens partial cDNA sequence; cl... 1593 4.0e-124
emb|F04964|HSC2OD122 H. sapiens partial cDNA sequence; cl... 1405 2.1e-108
gb|R41937|R41937 yg04g10.s1 Homo sapiens cDNA clone 3... 1083 1.4e-97 4
emb|F00058|HSB09H061 H. sapiens partial cDNA sequence; cl... 923 8.4e-87 3
[In the alignments click on to return to the respective one line description, or (if available) on to extract the entry from the database.]
gb|U07418|HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds.
Length = 2503
Plus Strand HSPs:
Score = 12515 (3458.1 bits), Expect = 0.0, P = 0.0
Identities = 2503/2503 (100%), Positives = 2503/2503 (100%), Strand = Plus / Plus
Query: 1 GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG 60
Sbjct: 1 GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG 60
Query: 61 TTATTCGGCGGCTGGACGAGACAGTGGTGAACCGCATCGCGGCGGGGGGAAGTTATCCAGC 120
Sbjct: 61 TTATTCGGCGGCTGGACGAGACAGTGGTGAACCGCATCGCGGCGGGGGGAAGTTATCCAGC 120
Query: 121 GGCCAGCTAATGCTATCAAAGAGATGATTGAGAACTGTTTAGATGCAAAATCCACAAGTA 180
Sbjct: 121 GGCCAGCTAATGCTATCAAAGAGATGATTGAGAACTGTTTAGATGCAAAATCCACAAGTA 180
Query: 181 TTCAAGTGATTGTTAAAGAGGGAGGCCTGAAGTTGATTCAGATCCAAGACAATGGCACCG 240
Document: Done.

Figure 2.5.8 The WWW interface of the BLAST network server (see Example 4): BLAST Output page.

Using BLAST Electronic Mail Server

```
------------------------------------------------------------------
mail blast@ncbi.nlm.nih.gov
Subject:

PROGRAM tblastn
DATALIB dbest
BEGIN
>HSCHM_1 choroideremia, Rab GGTase component A (REP-1) [Homo sapiens]
MADTLPSEFDVIVIGTGLPESIIAAACSRSGRRVLHVDSRSYYGGNWASFSFSGLLSWLK
```

Lines omitted for brevity.

```
VEKPRILWALYFNMRDSSDISRSCYNDLPSNVYVCSGPDCGLGNDNAVKQAETLFQEICP
NEDFCPPPPNPEDIILDGDSLQPEASESSAIPEANSETFKESTNLGNLEESSE
------------------------------------------------------------------
```

BOLASI / CENTRE UNIVERSITAIRE LIBRARY

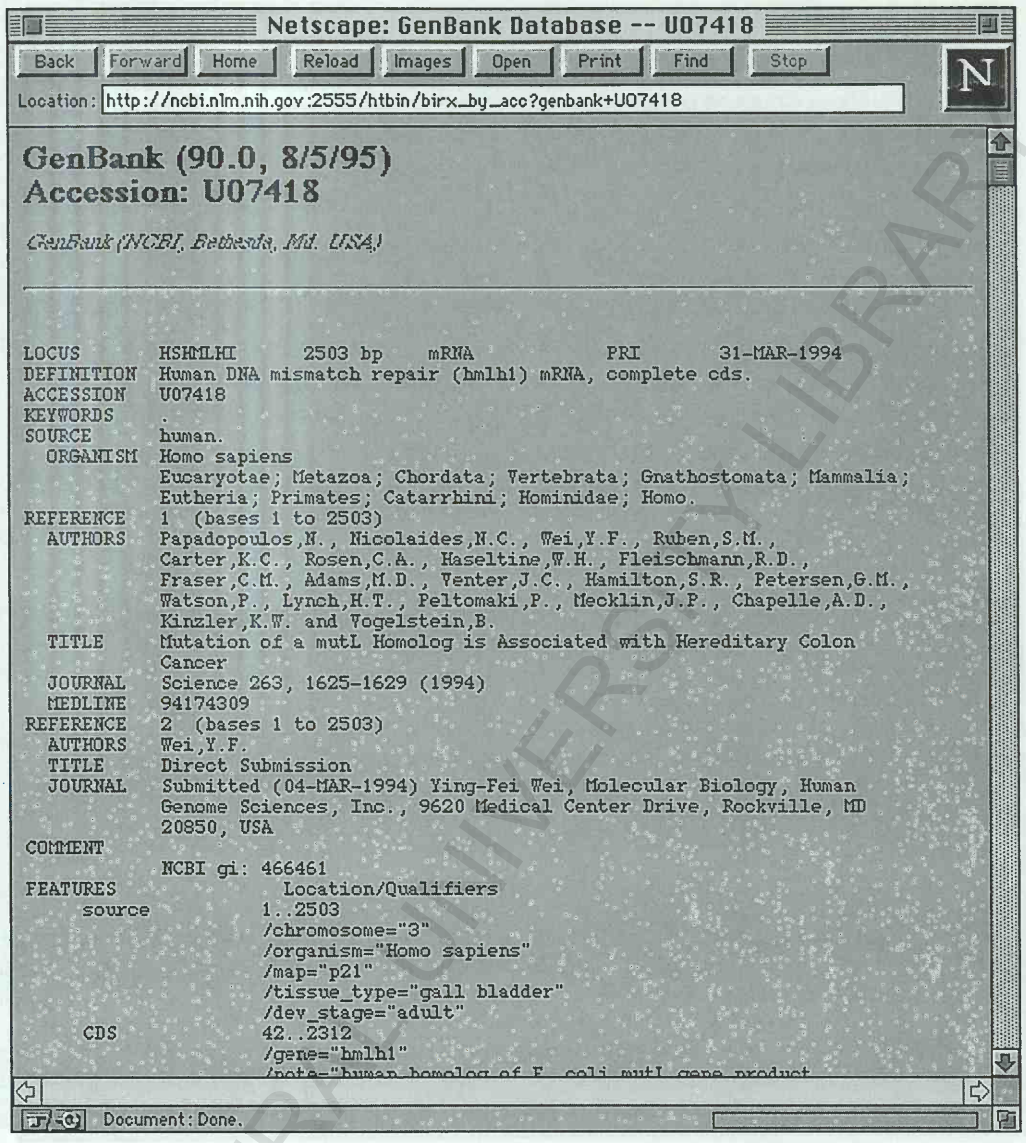


Figure 2.5.9 The WWW interface of the BLAST network server (see Example 4): Retrieve page.

Using BLAST Command-Line Client Program

The command-line equivalent of the above TBLASTN search is:

```
-------------------------------------------------------------------
my_computer> tblastn dbest p95.aa > p95_dbest.tblastn &
-------------------------------------------------------------------
```

Sample 3: Output from Example 5 Search

```
Query=  gi|460795|gp|X78121|HSCHM_1 choroideremia, Rab ger...
        (653 letters)

Database:  Database of Expressed Sequence Tags daily update (Mar 24)
           154,870 sequences; 50,913,880 total letters.
Searching.....................................................done
```


BIU JASLOK CENTRAL UNIVERSITY LIBRARY

```
                                                                    Smallest
                                                                      Sum
                                                    Reading  High  Probability
Sequences producing High-scoring Segment Pairs:       Frame  Score P(N)       N

gnl|dbest|5476   Z14454 cDNA Chris Martin sorted cDNA ... +1   118  8.0e-19  2
gnl|dbest|81031  T30998 cDNA Human Brain H.sapiens  Ho... +1   119  6.3e-18  2
gnl|dbest|38355  D25063 cDNA Rice root O.sativa  Homol... +1   122  8.9e-13  2
gnl|dbest|83003  T32970 cDNA Human Brain H.sapiens  Ho... +3   125  2.0e-12  2
gnl|dbest|59516  D35905 cDNA Yuji Kohara unpublished c... +1   134  3.4e-10  1
gnl|dbest|148763 F11715 cDNA normalized infant brain c... +3   126  6.0e-10  1
gnl|dbest|83691  T33658 cDNA Human Brain H.sapiens  Ho... +3   125  1.1e-09  1
gnl|dbest|81231  T31198 cDNA Human Brain H.sapiens  Ho... +3   124  1.4e-09  1
gnl|dbest|70908  D39245 cDNA Rice root O.sativa  Homol... +1   111  6.0e-08  2
gnl|dbest|5993   M89004 cDNA Chris Martin sorted cDNA ... +1   106  1.2e-07  2
gnl|dbest|121550 F05950 cDNA normalized infant brain c... +3   109  1.8e-07  1
gnl|dbest|28311  T09257 cDNA Infant Brain, Bento Soare... +3   117  2.5e-07  1
gnl|dbest|59895  D36307 cDNA Yuji Kohara unpublished c... +2    88  9.3e-06  2
gnl|dbest|46854  Z33929 cDNA Versailles-VB A.thaliana ... +2   104  2.6e-05  1
gnl|dbest|93102  T41840 cDNA Lambda-PRL2 A.thaliana  H... +1    83  0.0028   2
gnl|dbest|75937  D41864 cDNA Rice shoot O.sativa  Homo... +1    52  0.59     2
gnl|dbest|107612 T54329 cDNA Stratagene placenta (#937... -1    58  0.98     2
gnl|dbest|46480  T20253 cDNA Heart H.sapiens             +3    43  0.99     3
gnl|dbest|24978  T05924 cDNA Fetal brain, Stratagene (... +1    66  0.999    1
gnl|dbest|33207  Z26677 cDNA AC16H A.thaliana  Homolog... -3    46  0.99993  2

>gnl|dbest|5476 Z14454 cDNA Chris Martin sorted cDNA library C.elegans
           Homology: gp|U00002|CELRABGDI_1 rabGDI [Caenorhabditis elegans]
           Score: 552  pVal: 8.3e-72
           Length = 388

  Plus Strand HSPs:

 Score = 118 (54.1 bits), Expect = 8.0e-19, Sum P(2) = 8.0e-19
 Identities = 22/43 (51%), Positives = 30/43 (69%), Frame = +1

Query:     9 FDVIVIGTGLPESIIAAACSRSGRRVLHVDSRSYYGGNWASFS 51
             +D I +GTGL E II+   S SG++VLH+D  +YYGG  AS +
Sbjct:    55 YDAIXLGTGLKECIISGXLSVSGKKVLHIDRNNYYGGESASLT 183

 Score = 93 (42.6 bits), Expect = 8.0e-19, Sum P(2) = 8.0e-19
 Identities = 19/46 (41%), Positives = 27/46 (58%), Frame = +1

Query:   224 GRRFNIDLVSKLLYSRGLLIDLLIKSNVSRYAEFKNITRILAFREG 269
             GR  N+DL+ K   + G L+ LLI + V+RY EFK+I      + G
Sbjct:   250 GRXXNVDLIPKFXMANGPLVKLLIHTGVTRYLEFKSIEASFVVKGG 387

>gnl|dbest|81031 T30998 cDNA Human Brain H.sapiens  Homology: pir|S35964|S35964
           GDP dissociation inhibitor alpha - rat
           >gi|396431|gp|X74402|RNRABGDIA_1 rab GDI alpha [Rattus norvegicus]
           Score: 528  pVal: 2e-67
           Length = 326

  Plus Strand HSPs:

 Score = 119 (54.5 bits), Expect = 6.3e-18, Sum P(2) = 6.3e-18
 Identities = 22/44 (50%), Positives = 32/44 (72%), Frame = +1

Query:     8 EFDVIVIGTGLPESIIAAACSRSGRRVLHVDSRSYYGGNWASFS 51
             E+DVIV+GTGL E I++   S +G++VLH+D   YYGG  +S +
Sbjct:     4 EYDVIVLGTGLTECILSGIMSVNGKKVLHMDRNPYYGGESSSIT 135

 Score = 86 (39.4 bits), Expect = 6.3e-18, Sum P(2) = 6.3e-18
 Identities = 15/35 (42%), Positives = 25/35 (71%), Frame = +1
```

BOUAS / ETRA UNIVERSITY LIBRARY

```
Query:    224 GRRFNIDLVSKLLYSRGLLIDLLIKSNVSRYAEFK 258
              GR +N+DL+ K L + G L+ +L+ + V+RY +FK
Sbjct:    199 GRDWNVDLIPKFLMANGQLVKMLLYTEVTRYLDFK 303

>gnl|dbest|38355 D25063 cDNA Rice root O.sativa  Homology: sp|P21856|GDIS_BOVIN
            GDP DISSOCIATION INHIBITOR FOR SMG P25A (SMG P25A GDI).
            >pir|A35652|A35652 smg p25A regulatory protein - bovine
            >gp|D90103|BOVGDI_1 Bovine mRNA for regulatory protein for smg
            p25A. [Bos taurus] Score: 200  pVal: 3.3e-22
            Length = 357

  Plus Strand HSPs:

 Score = 122 (55.9 bits), Expect = 8.9e-13, Sum P(2) = 8.9e-13
 Identities = 24/54 (44%), Positives = 34/54 (62%), Frame = +1

Query:      2 ADTLPSEFDVIVIGTGLPESIIAAACSRSGRRVLHVDSRSYYGGNWASFSFSGL 55
              A  +  E+DVIV+GTGL E I++   S  G +VLH+D   YYGG+  S + + L
Sbjct:     64 AGXMDEEYDVIVLGTGLKECILSGLLSVDGLKVLHMDRNDYYGGDSTSLNLNQL 225

 Score = 45 (20.6 bits), Expect = 8.9e-13, Sum P(2) = 8.9e-13
 Identities = 7/19 (36%), Positives = 13/19 (68%), Frame = +1

Query:    225 RRFNIDLVSKLLYSRGLLI 243
              R +N+D+V K + + G L+
Sbjct:    277 RDYNVDMVPKFMMANGTLV 333
```

Remainder of the BLAST output omitted for brevity.

Discussion

The results of searching the human choroideremia protein query against dbEST are shown in Sample 3 (the reading frame in this case is that of the database sequence). Note that there are fourteen significant hits. There are seven matches from human proteins, four matches to nematode proteins, and three matches to plant sequences (rice and *Arabidopsis*). Given the vast evolutionary distance between mammals, nematodes, and plants, it is doubtful that these sequence similarities would be detectable on the nucleotide level and indeed they are not detected if one uses BLASTN to search dbEST with the choroideremia cDNA sequence (not shown).

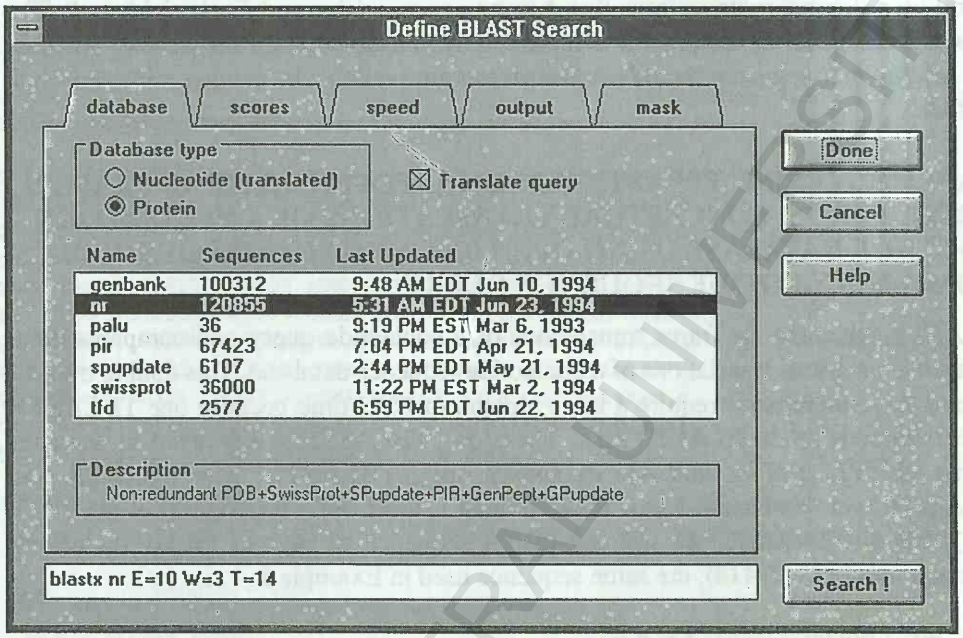
As in the BLASTX example (Sample 1), a reading frame is included in both the hit list and the longer descriptions associated with the alignments. In this example, however, the reading frame indicates the reading frame of the database sequence, not the query sequence, which is an amino acid sequence.

BIU IASI / CENTRAL UNIVERSITY LIBRARY

EXAMPLE 6: USING THE NetBLAST GRAPHICAL CLIENT PROGRAM

Here we illustrate the new NetBLAST graphical interface by repeating the search described in Example 2 but with SEG + XNU masking that makes the output more similar to that of Example 3. Figure 2.5.10A shows the dialog box used to set up a search request and Figure 2.5.10B shows one view of the search results where the alignments are depicted schematically. Compare this view with the textual version of the alignments shown in Sample 2. The NetBLAST interface provides a global view of local alignments.

The various choices of databases, parameters, and options are grouped into several categories, each of which is represented by a "folder tab" across the top of the dialog box in the NetBLAST graphical client program (Fig. 2.5.10A). In this example, the Database tab has been selected. Radio buttons are used to toggle between nucleotide and protein database types and the current list of databases of the selected type is shown in the list



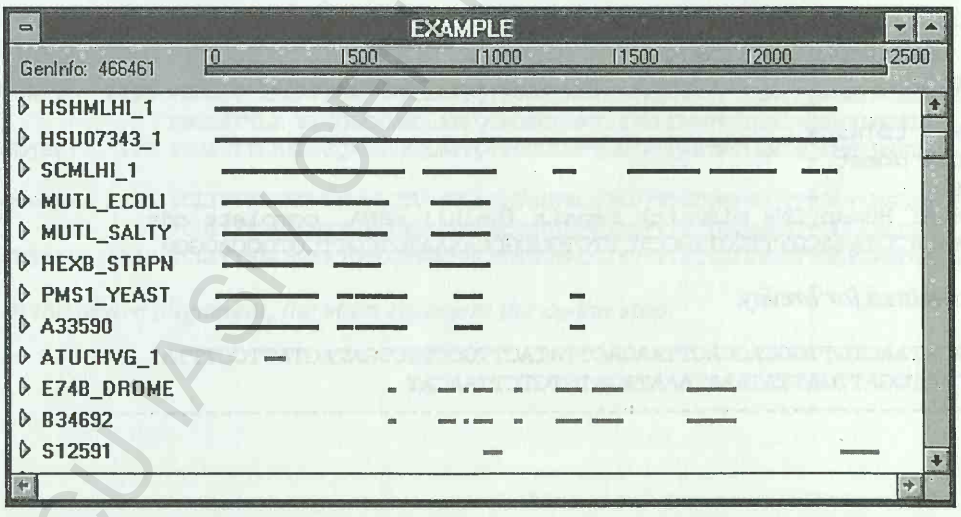


Figure 2.5.10 The NetBLAST graphical client program (see Example 4) (**A**) Dialog box. (**B**) Hit list from a BLASTX search with human MHL1 query sequence using the NetBLAST graphical client program.

FCU IASW/CENTRAL UNIVERSITY LIBRARY

box. The desired BLAST command can be determined from the type of query sequence, whether or not it is to be translated (Translate Query checkbox, which is enabled for nucleotide sequences only), and the type of database. In this example, the nucleotide query sequence is to be translated and compared to a protein database, which implies that the BLASTX command should be used, as shown in the field at the bottom of the window.

The hit list from a BLASTX search with human MLH1 query sequence using the NetBLAST graphical client program is displayed in a window (Fig. 2.5.10B). A scale across the top shows the extent of the query sequence (measured in bases in this example, because the query is a nucleotide sequence). Identifiers of protein sequences bearing similarities to the translated query are listed below, along with a graphical indication of the locations of the HSPs within the query sequence. The first two "self hits" arise from the CDS translations of the GenBank entries for the human DNA mismatch repair cDNA (two versions). The short regions at the ends of the query sequence that match nothing in these proteins are the 5′ and 3′ untranslated regions. Below the human sequences are several homologous proteins from *Saccharomyces cerevisiae* and *Escherichia coli*. The amino-terminal half of the protein is more highly conserved than the carboxy-terminal half. Alignments, sequence records, and abstracts are available by using a point-and-click interface.

EXAMPLE 7: USING TBLASTX TO SEARCH OPEN READING FRAMES DERIVED FROM CONCEPTUAL TRANSLATIONS OF A NUCLEOTIDE SEQUENCE DATABASE USING OPEN READING FRAMES DERIVED FROM A NUCLEOTIDE SEQUENCE QUERY

TBLASTX takes the six-frame translations of a nucleotide query and compares them against the six-frame translations of a nucleotide sequence database. This double translation and the comparisons require a large amount of CPU time because one TBLASTX run corresponds to 36 BLASTP runs. Therefore, TBLASTX is restricted to searches against dbEST, dbSTS, and *Alu* databases only. TBLASTX, like BLASTX and TBLASTN, is a fundamental tool for searching against dbEST (Boguski et al., 1993, 1994). For this example, we use the cDNA sequence of human MLH1 (GenBank accession number U07418), the same sequence used in Example 1.

Using BLAST Electronic Mail Server

```
-----------------------------------------------------------------
mail blast@ncbi.nlm.nih.gov
Subject:

PROGRAM tblastx
DATALIB dbest
BEGIN
>HSHMLHI Human DNA mismatch repair (hmlh1) mRNA, complete cds
GTTGAACATCTAGACGTTTCCTTGGCTCTTCTGGCGCCAAAATGTCGTTCGTGGCAGGGG
```

Lines omitted for brevity.

```
ACCAACATAAGTGTTGGTAGCACTTAAGACTTATACTTGCCTTCTGATAGTATTCCTTTA
TACACAGTGGATTGATTATAAATAAATAGATGTGTCTTAACAT
-----------------------------------------------------------------
```


RAJIV GANDHI UNIVERSITY LIBRARY

Using BLAST Command-Line Client Program

The command line equivalent of the above BLASTN search is:

```
------------------------------------------------------------------
my_computer> tblastx dbest hmlh1.seq > hmlh1_genbank.tblastx
------------------------------------------------------------------
```

Sample 4: Output from Example 7 Search

```
Query=  gi|466461|gb|U07418|HSHMLHI Human DNA mismatch repair (hmlh1) mRNA,
                          complete cds.
        (2503 letters)

  Translating both strands of query sequence in all 6 reading frames

Database:  Database of Expressed Sequence Tags daily update (May 12)
           207,556 sequences; 69,849,438 total letters.
Searching.....................................................done

                                                                  Smallest
                                                                    Sum
                                                  Reading High  Probability
Sequences producing High-scoring Segment Pairs:     Frame Score P(N)       N

gnl|dbest|78657  T28624 cDNA Human Gall bladder H.sapi. . +1  731  2.3e-93  1
gnl|dbest|29995  Z24775 cDNA STRATAGENE Human skeletal. . +1  577  9.7e-72  1
gnl|dbest|120316 F04253 cDNA normalized infant brain c. . +1  564  6.5e-70  1
gnl|dbest|120340 F04964 cDNA normalized infant brain c. . +1  504  1.7e-61  1
gnl|dbest|149607 F00058 cDNA STRATAGENE Human skeletal. . -1  332  1.3e-46  2
gnl|dbest|212002 R41937 cDNA Soares infant brain 1NIB . . +2  393  3.6e-46  1
gnl|dbest|180958 R10662 cDNA Soares fetal liver spleen. . +3  348  9.3e-40  1
gnl|dbest|59205  D37616 cDNA Yuji Kohara unpublished c. . +3  174  9.7e-16  1
gnl|dbest|60361  D36773 cDNA Yuji Kohara unpublished c. . +3  168  7.2e-15  1
gnl|dbest|88910  T38294 cDNA S. cerevisiae strain X218. . +1  124  3.6e-07  1
gnl|dbest|59227  D37638 cDNA Yuji Kohara unpublished c. . +3   92  0.014    1

>gnl|dbest|78657 T28624 cDNA Human Gall bladder H.sapiens  Homology:
          gi|466462|gp|U07418|HSHMLHI_1 hmlh1 gene product [Homo sapiens]
          Score: 363  pVal: 2e-43
          Length = 438

  Minus Strand HSPs:

 Score = 704 (323.8 bits), Expect = 1.4e-89, P = 1.4e-89
 Identities = 135/136 (99%), Positives = 135/136 (99%), Frame = +1 / -3

Query:  2095 PSGSSTYLRSRPSQASRVKCLAPFQTPGSGLWNTLSIKPCAHTFCLLNISQKMEISCSLL 2274
              PSGSSTYLR RPSQASRVKCLAPFQTPGSGLWNTLSIKPCAHTFCLLNISQKMEISCSLL
Sbjct:   409 PSGSSTYLRXRPSQASRVKCLAPFQTPGSGLWNTLSIKPCAHTFCLLNISQKMEISCSLL 230

Query:  2275 TCLIYTKSLRGVKYGYLCTVGCVLLSLYSDTKCCIKV*YTKCTNISVGST*DLYLPSDSI 2454
              TCLIYTKSLRGVKYGYLCTVGCVLLSLYSDTKCCIKV*YTKCTNISVGST*DLYLPSDSI
Sbjct:   229 TCLIYTKSLRGVKYGYLCTVGCVLLSLYSDTKCCIKV*YTKCTNISVGST*DLYLPSDSI 50
```

In the above alignment, the stars represent the codon stop.

BCU IASI CENTRAL UNIVERSITY LIBRARY

```
Query:  2455 PLYTVD*L*INRCVLT 2502
             PLYTVD*L*INRCVLT
Sbjct:    49 PLYTVD*L*INRCVLT 2

  Plus Strand HSPs:

 Score = 731 (336.3 bits), Expect = 2.3e-93, P = 2.3e-93
 Identities = 134/136 (98%), Positives = 134/136 (98%), Frame = -1 / +1

Query:  2503 MLRHIYLFIINPLCIKEYYQKASISLKCYQHLCWYTLYITL*YNTLYRNTEKEEHIPQCI 2324
             MLRHIYLFIINPLCIKEYYQKASISLKCYQHLCWYTLYITL*YNTLYRNTEKEEHIPQCI
Sbjct:     1 MLRHIYLFIINPLCIKEYYQKASISLKCYQHLCWYTLYITL*YNTLYRNTEKEEHIPQCI 180

Query:  2323 NNHI*HLSKTLYRSGRLASCRIFPSSVKCLGGRMCERKAL*TMCSTVHFQEFGMEPGTSL 2144
             NNHI*HLSKTLYRSGRLASCRIFPSSVKCLGGRMCERKAL*TMCSTVHF EFGMEPGTSL
Sbjct:   181 NNHI*HLSKTLYRSGRLASCRIFPSSVKCLGGRMCERKAL*TMCSTVHFXEFGMEPGTSL 360

Query:  2143 CWPERVDSSDMYCFRM 2096
             CWPERV SSDMYCFRM
Sbjct:   361 CWPERVXSSDMYCFRM 408

 Score = 700 (322.0 bits), Expect = 4.9e-89, P = 4.9e-89
 Identities = 136/136 (100%), Positives = 136/136 (100%), Frame = -3 / +3

Query:  2501 VKTHLFIYNQSTVYKGILSEGKYKS*VLPTLMLVHFVYHTLIQHFVSEYRERRTHPTVHK 2322
             VKTHLFIYNQSTVYKGILSEGKYKS*VLPTLMLVHFVYHTLIQHFVSEYRERRTHPTVHK
Sbjct:     3 VKTHLFIYNQSTVYKGILSEGKYKS*VLPTLMLVHFVYHTLIQHFVSEYRERRTHPTVHK 182

Query:  2321 *PYLTPLKDFV*IRQVSKLQDISIFCEMFRRQNV*AQGFIDNVFHSPLPGVWNGARHFTL 2142
             *PYLTPLKDFV*IRQVSKLQDISIFCEMFRRQNV*AQGFIDNVFHSPLPGVWNGARHFTL
Sbjct:   183 *PYLTPLKDFV*IRQVSKLQDISIFCEMFRRQNV*AQGFIDNVFHSPLPGVWNGARHFTL 362

Query:  2141 LA*EGRLLRYVLLPDG 2094
             LA*EGRLLRYVLLPDG
Sbjct:   363 LA*EGRLLRYVLLPDG 410

 Score = 673 (309.6 bits), Expect = 2.9e-85, P = 2.9e-85
 Identities = 134/136 (98%), Positives = 134/136 (98%), Frame = -2 / +2

Query:  2502 C*DTSIYL*SIHCV*RNTIRRQV*VLSATNTYVGTLCISHFDTTLCIGIQRKKNTSHSA* 2323
             C*DTSIYL*SIHCV*RNTIRRQV*VLSATNTYVGTLCISHFDTTLCIGIQRKKNTSHSA*
Sbjct:     2 C*DTSIYL*SIHCV*RNTIRRQV*VLSATNTYVGTLCISHFDTTLCIGIQRKKNTSHSA* 181

Query:  2322 ITIFNTSQRLCIDQAG*QAAGYFHLL*NV*EAECVSARLYRQCVPQSTSRSLEWSQALHS 2143
             ITIFNTSQRLCIDQAG*QAAGYFHLL*NV*EAECVSARLYRQCVPQSTS SLEWSQALHS
Sbjct:   182 ITIFNTSQRLCIDQAG*QAAGYFHLL*NV*EAECVSARLYRQCVPQSTSXSLEWSQALHS 361

Query:  2142 AGLRGSTPQICTASGW 2095
             AGLRGS PQICTASGW
Sbjct:   362 AGLRGSXPQICTASGW 409
```

Remainder of the BLAST output omitted for brevity.

Discussion

The results of the cDNA human hmlh1 sequence search against dbEST are shown in Sample 4. The hit list and the alignments from the first hits (dbEST-78657) are given. Two reading frames are indicated for each alignment, one for the database sequence and one for the query sequence (Frame = database/query). However, only the database sequence is listed in the hit list. In this example, the alignments correspond to four different searches using four different protein query sequences against four different protein databases. BLOSUM62 is used as the default scoring scheme but this selection can be changed easily (see manual and e-mail help files).

BU IAS/CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

The earliest sequence comparison studies attempted to produce end-to-end alignments of complete sequences, also known as global alignments. This was because the sequences then available for study were those of small, globular proteins consisting of a single domain, such as cytochromes or hemoglobin subunits. Also, these alignment techniques were developed before the discovery of the exon/intron structure of genes. However, it is now recognized that most genes and proteins are modular and mosaic structures often consisting of several distinct domains and/or repeated copies of domains, the order of which may be rearranged by several genetic mechanisms (Patthy, 1991; Doolittle and Bork, 1993). Therefore, attention has shifted to algorithms for local alignment, and virtually all database search strategies in use today seek local alignments. See Altschul et al. (1994) for additional background material and references on alignment algorithms.

The classic method for finding optimal local alignments was developed by Smith and Waterman (1981). Given the large sizes of sequence databases today, it is impractical to use this algorithm except on special massively parallel computers (Brutlag, 1993). MPsrch is one such parallel Smith-Waterman implementation that is available via a public e-mail server called Blitz (contact netserv@ebi.ac.uk for more information). MPsrch is also available commercially from Intelligenetics. In addition to being slow to compute, another drawback of this algorithm is that no rigorous theory is available to evaluate the statistical significance of the gapped alignments produced (Altschul et al., 1994).

"Heuristic," as opposed to "optimal," database search methods are obtained by implementing certain shortcuts—for example, forbidding the introduction of gaps (insertion/deletion mutations) in local alignments, at least during the initial phase of the search. Although in principle heuristic methods may be less sensitive for detecting weak similarities, they are much faster when run on general-purpose computers. The two most popular heuristic search programs in use today are FASTA (Pearson and Lipman, 1988; Pearson, 1990) and BLAST (Altschul et al., 1990). The following discussion focuses on BLAST, beginning by comparing and contrasting it to FASTA.

The similarity in approach between FASTA and BLAST lies in their initial steps—both programs begin by identifying ungapped aligned segments, although they use different heuristics. In this regard, specific implementations of the BLAST programs provide the most leeway in adjusting the speed of the search, albeit at the expense of sensitivity at higher speeds. The biggest difference between the two approaches arises in what the programs do with this initial alignment information. FASTA picks the highest-scoring, ungapped aligned segment it has found and performs a local (Smith-Waterman) alignment confined within a narrow window centered about that segment. Gaps may be introduced in generating the optimized alignment. FASTA then ranks the database sequences by these optimized scores. Statistical significance of FASTA and Smith-Waterman scores is often estimated by relating individual scores to the mean and standard deviation of the scores obtained from the entire database search. However, assuming a Gaussian distribution greatly overestimates the significance (i.e., underestimates the probability) of these optimized alignment scores. Optimized scores tend to follow an extreme-value distribution which falls off by e^{-x}, where x is proportional to the alignment score, rather than the rapid e^{-x^2} decay of a Gaussian distribution.

In contrast, once the BLAST programs have found some aligned segments (using the BLAST algorithm), the segments are considered in their totality for their statistical significance. For example, two regions of similarity that are moderately to widely

BOSTON UNIVERSITY LIBRARY

separated from one another (e.g., two conserved domains separated by variable-length intervening regions having little or no discernible similarity) will be considered together by the BLAST programs and may be found to achieve statistical significance, whereas an optimized gapped alignment centered around any one of the two regions might not yield a statistically significant score.

The limitation of using ungapped alignments in BLAST programs has the benefit of making estimation of their statistical significance a tractable problem. For this purpose, Karlin-Altschul (K-A) statistics are used (Karlin and Altschul, 1990, 1993). K-A statistics afford a rapid, analytical solution to the statistics problem and are often quite accurate. With more accurate statistical significance estimates, it is possible to set a more stringent threshold for reporting database matches, thus reducing the number of random matches that must be manually considered and probably rejected. It must be noted that this is an absolute necessity for high-throughput sequencing projects that must manage data at rates of 1 megabase or more per week.

With real-world sequence data (not random models of sequence data), K-A statistics tend to fall down most frequently and most dramatically when two sequences being compared contain regions of statistically biased residue composition (e.g., simple sequence repeats or SSRs). To address this problem, the BLAST family of programs support a "filter" (filter) option to have these compositionally biased regions masked from the query sequence prior to searching. The principles and practice of sequence masking are discussed in more detail below.

The BLASTN program for comparing a nucleotide query sequence against a nucleotide database is insensitive at identifying even moderately distant homologs and is not intended to be used for this purpose. FASTA is far better at this application, but not nearly so fast as BLASTN. Most of these differences in speed and sensitivity arise from BLASTN's current restriction to using a word length (see BLAST Search Algorithm) of 12, while FASTA can use a word length (or "ktup") of length 1 to 6. What this means is that there must be at least twelve consecutive matching nucleotides for BLASTN to key in on a similar segment, whereas FASTA needs as little as one matching nucleotide. The larger this initial word length, the faster each program runs, but the sensitivity suffers at the same time.

With BLASTN, however, the sensitivity problem is often not an issue because better sensitivity in identifying similarity between coding regions can be obtained by comparing the translation products rather than the encoded nucleotide sequences (see Examples and States et al., 1991). This arises from the redundancy of the genetic code, the small size of the DNA alphabet compared with that of proteins, and the fact that it is possible to take advantage of positively scoring imperfect matches (conservative amino acid substitutions) for proteins but not nucleic acids. In searching for similarity between nucleotide regulatory regions, however, FASTA remains superior.

Some final distinctions between BLAST and FASTA include the following: BLASTN, BLASTX, TBLASTN, and TBLASTX compare both strands of the query by default, whereas some implementations of FASTA do not; for protein-protein comparisons using default parameter values, BLASTP rivals or exceeds FASTA in sensitivity and specificity and runs considerably faster; there is no functional equivalent in FASTA to BLASTX, TBLASTN, or TBLASTX; FASTA runs on more computers than BLAST; and FASTA reads more database formats than do the BLAST programs. This last apparent advantage is nullified by the availability of comprehensive nonredundant databases that are updated daily and available for searching by BLAST e-mail or network services. FASTA e-mail

servers are supported at a number of sites (see *UNIT 2.4*). FASTA source code is available from Dr. William Pearson by anonymous FTP (see *UNIT 2.4*).

Scoring Systems and Alignment Statistics

In order to rank the results of a database search, it is necessary to have some way to calculate the quality of each individual alignment. This is done using a scoring system, which normally consists of a table of residue substitution scores and rules for its use in calculating an alignment score. The default scoring systems used by the BLAST family of programs are BLOSUM62 (Henikoff and Henikoff, 1992) when comparing proteins and a simple unitary matrix when comparing nucleotide sequences. Although these scores work well in typical cases, there may be reasons to select an alternate scoring system using the M= (command-line) or MATRIX (e-mail) options (more than 65 different matrices are available). Scores that are best for detecting distantly related sequences differ from those suited to finding nearly identical short motifs (Altschul, 1991, 1993). Special scoring systems have also been developed for detecting frameshift sequencing errors (Claverie, 1993). The latest version of BLAST (version 1.4) supports a multiple matrix option.

The BLAST search algorithm seeks a type of local alignment called a high-scoring segment pair (HSP). An HSP is a pair of sequence segments, one drawn from the query sequence and the other from the database sequence, that may be aligned without gaps to achieve a high score. Calculation of the score of an HSP is a simple matter of summing the individual substitution scores for each aligned pair of letters. The end points of an HSP are chosen such that the score is locally maximal. In order to be reported by BLAST, the score must meet or exceed a score threshold (S). It is possible to specify the value of S as an input parameter.

Given an alignment reported by a database search, an important question is whether it represents a true relationship or is due simply to chance. Given certain characteristics of the query sequence, the database, and the scoring system, mathematical analysis can be used to calculate a probability density function for the distribution of HSP scores resulting from a database search (Karlin and Altschul, 1990). To see if a particular database hit should be considered significant, simply compare its HSP score to this distribution. More precisely, calculate a p value, which gives the probability that an HSP with a score as good as the one observed could be due simply to chance.

Because the HSP score that must be achieved to be considered significant depends on certain aspects of the search context, especially the scoring system, it is difficult to know the appropriate value of S to use in a search. For this reason, it is more convenient to specify the sensitivity of the search in terms of the significance threshold (E), which takes the search context into account. The E parameter may be thought of as the number of matches expected to be observed by chance alone during the database search. Given the statistical model of Karlin and Altschul, the values of E and S are interconvertible—based only on knowledge of the scoring system, the length and composition of the query sequence, and the size and hypothetical composition of the database.

BLAST Search Algorithm

Having defined what is needed in a database search, the next question is how to perform the search for HSPs in a reasonable period of time. When comparing two sequences of length N, an exhaustive search for HSPs would require time proportional to N^2. Such a time requirement would be prohibitive when searching a sequence database containing hundreds of thousands of sequences. To get around this problem, many rapid search algorithms make use of word-based methods. Imagine that the residues in the query sequence are taken several at a time instead of singly. These groups of letters are known

BIOLIST CITRA UNIVERSITY LIBRARY

as words and the number of letters is called the word size (W). Related sequences are likely to share many common words, whereas unrelated sequences will have fewer. By limiting the search for similarities to only the regions near matching words, the search space that must be examined is dramatically reduced, leading to faster search times. The time required for the search is still proportional to N^2, but the proportionality constant is small.

Although the use of word-based methods can lead to faster search times, these algorithms are heuristic rather than rigorous and there is some chance that weaker, yet still statistically significant similarities may be missed. This is because the sequences may have diverged to the point where no common words of length W remain. This tradeoff between speed and sensitivity is affected by the value chosen for W. Larger W leads to faster searches while sacrificing sensitivity. On the other hand, smaller W leads to slower but more thorough searches.

Most word-based methods require exact matching of the letters within the word. Unfortunately, under these conditions obtaining satisfactory sensitivity may require reducing W to such a small value (perhaps 2, or even 1) that the speed advantage afforded by using a word-based method is nearly negated. BLAST introduces a novel method to deal effectively with this problem. Essentially, the same score matrix used to calculate the quality of an HSP is applied to the generation of words from the query sequence. For each word in the query, a list of all words that would produce a score of at least T using the score matrix is compiled. This extended group of words is called the neighborhood and T is the neighborhood score threshold. The concept of neighborhood words is illustrated in Figure 2.5.11. The value of T, like that of W, affects the speed and sensitivity of the search. Increasing the value of T results in faster searches because there will be fewer neighborhood words to check for in the database sequence.

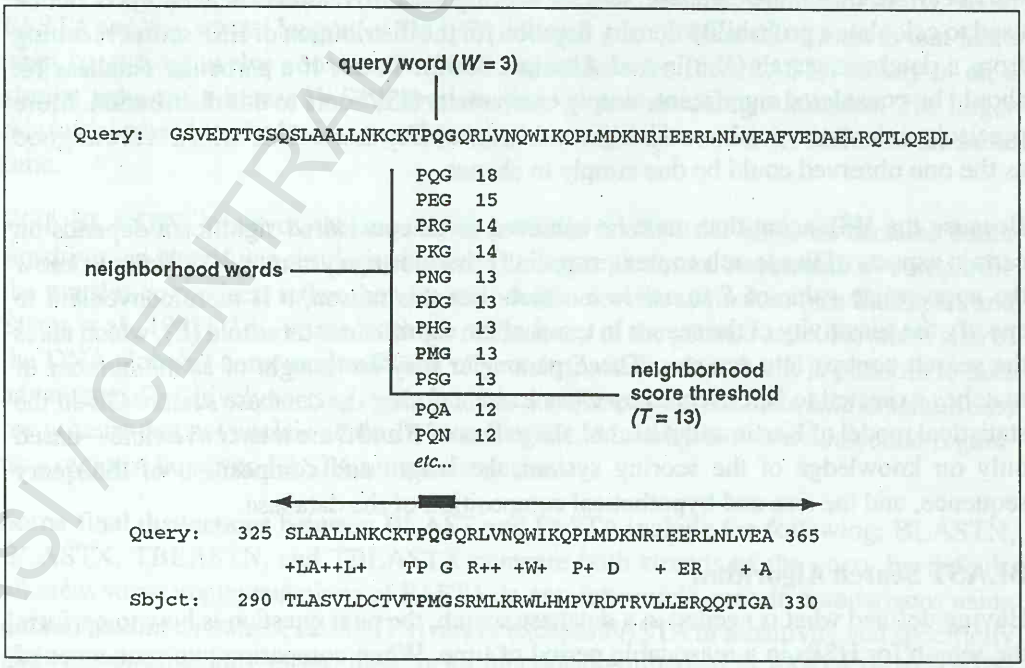


Figure 2.5.11 Neighborhood words used by the BLAST search algorithm. The BLAST search algorithm uses a word-based (or hashing) scheme to improve speed, but uses neighborhood words to preserve sensitivity. Shown here is an example of the neighborhood word list generated for a single word (PQG or Pro-Gln-Gly) in the query sequence using $W = 3$ and $T = 13$. Once a word match is encountered between the query and a database sequence, an attempt is made to extend the match in both directions to achieve a locally maximal score, the HSP score.

BU IASI / CENTRAL UNIVERSITY LIBRARY

Once the neighborhood words have been tabulated for the query sequence, each sequence in the database is scanned for the presence of matching words. For each word hit found in this way, an attempt is made to extend the match to the left and to the right (Fig. 2.5.8) to see if an HSP having a score of at least *S* can be identified. A typical approach is to continue extension until the cumulative score reaches zero and then take the point at which the maximum value of the cumulative score was achieved as the end of the run. However, BLAST introduces an additional heuristic in which the extension can also be terminated if the cumulative score drops off by some value *X* or more. For typical values of *X*, this speeds up the search without appreciably affecting the results. Reducing *X* will lead to further performance enhancement, but could result in missing some significant similarities. Increasing *X* to some very large value will ensure that no hits are missed as a result of premature hit termination.

Low-Complexity Sequences

The statistical framework upon which BLAST is based assumes that proteins can be modeled as random sequences. However, many real proteins contain patches of distinctly nonrandom sequence. These patches may be homopolymeric tracts, short-period repeats, or merely regions enriched in one or a few amino acids. Such low-complexity (LC) sequences are surprisingly abundant in real proteins; it has been observed that at least 25% of proteins in the database contain at least one LC region (Wootton, 1994a,b). Very little is known about the structure, function, or evolution of LC segments.

Despite the fact that LC sequences may well have interesting properties and functions, they wreak havoc on database searches. This is primarily because the residue frequencies within such regions differ markedly from the database as a whole. The result is that pairs of proteins containing LC sequences rich in the same amino acids may be aligned with each other to achieve high scores even though other regions of the proteins are entirely dissimilar. Because these alignments do not reflect a common ancestry, no functional inference is justified despite their high statistical significance. Even a small LC segment in a query sequence (e.g., see Example 3) can generate a multitude of these uninteresting hits, making the output difficult to interpret if not outright unmanageable. In some cases, the high volume of LC hits can completely swamp potentially interesting matches to distantly related proteins.

Two approaches have been devised to locate LC regions in query sequences. The SEG program (Wootton and Federhen, 1993) uses a rigorous mathematical definition of complexity and optimally segments a sequence into regions of alternating high and low complexity (Fig. 2.5.12). The XNU program (Claverie and States, 1993) specifically seeks short-period repeats by looking for runs of high score when the sequence is offset by short distances and compared to itself. The two methods are complementary: each can identify regions not found by the other, although the overall results are similar. Both programs support a mode of operation in which the sequence is written back out with the residues of the identified segments converted to Xs (an ambiguity code that means "any residue"). The strategy of "X-ing out" specific regions of the query sequence prior to a database search is known as "masking" or "filtering." For the BLASTP, BLASTX, and TBLASTX programs, the -filter command-line option (or the FILTER e-mail directive) followed by seg, xnu, or seg + xnu can be used to instruct the program to perform automatic masking on the query sequence using the algorithms separately or in combination.

BCU INFORMATION SERVICES CENTRAL UNIVERSITY LIBRARY

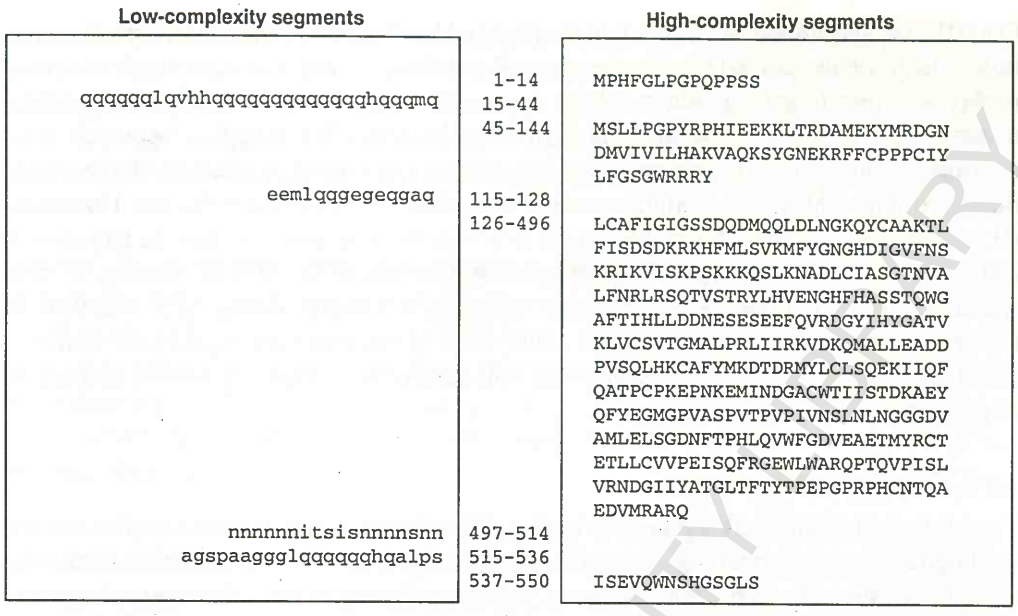


Figure 2.5.12 Low-complexity segments reported by the seg program. Subjecting the fruit fly RBP-J_κ sequence to analysis by seg produces the output shown here, in which the low-complexity segments are printed on the left in lower case and the high-complexity segments on the right in upper case. When this sequence is used as a query for a BLASTP search using the seg filter option, the letters of the low-complexity segments are converted to Xs prior to performing the database search.

STRATEGIES AND TIPS

Screening for Query Contamination

Imagine the frustrating results of screening a clone library with a probe that has been contaminated with cloning vector sequences. Or suppose that a probe used to hybridize against human genomic DNA contains a portion of an *Alu* repetitive element. Just as when performing a hybridization experiment, the quality of the query sequence that is used to "probe" for homology against a sequence database is critical to obtaining meaningful BLAST results. A newly-generated sequence that has been subjected to relatively little scrutiny may well be contaminated with vector sequences. Other sorts of contamination, for example with mitochondrial or bacterial sequences, are also possible. In addition, the presence of repetitive elements or structural RNA genes represents a sort of natural "contamination." In all of these cases, many hits will be produced that have a high statistical significance, but they will be frustrating and minimally informative, and may obscure more subtle but interesting homologies.

In some cases, the problem of query contamination can be difficult to recognize. Based on the analogy between database searching and molecular hybridization, it might be expected that a query sequence containing a repetitive element would generate extremely high numbers of matches, many of which would look essentially alike. The copy numbers for repetitive elements in sequence databases, however, are far smaller than in genomes, so the volume of search output may not be so high as to cause alarm. Furthermore, many of the copies that do exist are small portions of other gene sequences, and their presence is rarely mentioned in the definition lines in the hit list or alignment section. They may not even be annotated in the full database record describing a sequence. However, one thing to watch for when searching GenBank and Swiss-Prot is definition lines for any of several representative *Alu* repeat sequences that have been included as "warning entries." These warning entries are:

RAJIV GANDHI UNIVERSITY LIBRARY

GenBank:

```
U14567 ***ALU WARNING: Human Alu-J subfamily consensus sequence
U14568 ***ALU WARNING: Human Alu-Sb subfamily consensus sequence
U14569 ***ALU WARNING: Human Alu-Sb1 subfamily consensus sequence
U14570 ***ALU WARNING: Human Alu-Sb2 subfamily consensus sequence
U14571 ***ALU WARNING: Human Alu-Sc subfamily consensus sequence
U14572 ***ALU WARNING: Human Alu-Sp subfamily consensus sequence
U14573 ***ALU WARNING: Human Alu-Sq subfamily consensus sequence
U14574 ***ALU WARNING: Human Alu-Sx subfamily consensus sequence
```

Swiss-Prot:

```
P23959|ALUA_HUMAN !!!! ALU CLASS A WARNING ENTRY !!!!
P23960|ALUB_HUMAN !!!! ALU CLASS B WARNING ENTRY !!!!
P23961|ALUC_HUMAN !!!! ALU CLASS C WARNING ENTRY !!!!
P23962|ALUD_HUMAN !!!! ALU CLASS D WARNING ENTRY !!!!
P23963|ALUE_HUMAN !!!! ALU CLASS E WARNING ENTRY !!!!
P23964|ALUF_HUMAN !!!! ALU CLASS F WARNING ENTRY !!!!
```

It is preferable to detect query contamination by explicit test rather than chance discovery. One way to do this is to perform preliminary BLAST searches against special sequence databases containing various classes of problematic sequences. For example, the BLAST server provides access to: (1) a vector database which contains nucleotide sequences of a number of standard cloning vectors; (2) the alu amino acid sequence database which contains six-frame translations of representative human *Alu* repeats; and (3) the alu nucleotide sequence database which contains representative *Alu* elements (Jurka et al., 1992; Claverie and Makalowski, 1994). Preliminary screening of all new sequences for possible contamination is highly advisable.

Masking Low-Complexity Subsequences

As discussed in the section on Low-Complexity Sequences and illustrated in Example 3, the filter option in BLAST can be used along with seg, xnu, or seg + xnu commands to instruct the program to detect LC sequences in the query and exclude them from the search. Use of this option is highly recommended for all database searches that involve comparisons at the amino acid level. To illustrate the effect that LC masking can have on a database search, consider a BLASTP search using the translated product of the fruit fly RBP-J_κ gene (Furukawa et al., 1991) as the query sequence. The normal function of this gene is to encode a protein that specifically binds to the immunoglobulin J_κ recombination signal sequence.

When the RBP-J_κ sequence is used in a BLASTP search of the the nonredundant database without LC masking, 898 sequences are found to contain at least one HSP (by default only the best 250 are reported). The hit list from this search is shown in Figure 2.5.13. Several "self hits" involving other RBP-J_κ sequences in the database are observed (marked by a square in Fig. 2.5.13), followed by many weaker matches that are reported to be statistically significant. Many of the remaining sequences are false positives whose similarity is merely due to the fact that the segments contain unusually high proportions of glutamine (Q) and asparagine (N; see Fig. 2.5.12). With so many false positives, it is tempting to abandon further examination of the results, but a look at the 516th sequence in the list (marked by a star in Figs. 2.5.13 and 2.5.14) reveals an interesting similarity to the human fork head–related protein (Hromas et al, 1993). The alignment consists of a single high-scoring segment, shown in Figure 2.5.14, that does not fall within a low-complexity region of RBP-J_κ.

BOSTON UNIVERSITY LIBRARY

```
                                                                  Smallest
                                                                    Sum
                                                          High    Probability
  Sequences producing High-scoring Segment Pairs:         Score   P(N)        N

• pir|S15463|S15463       J-kappa recombination sequence-bin...  2942  0.0         1
• sp|P28159|RBJK_DROME    J KAPPA-RECOMBINATION SIGNAL BINDI...  2818  0.0         2
• sp|P31266|RBJK_MOUSE    J KAPPA-RECOMBINATION SIGNAL BINDI...  1898  3.7e-260    1
• sp|Q06330|RBJK_HUMAN    J KAPPA-RECOMBINATION SIGNAL BINDI...  1877  3.2e-257    1
• gp|M81870|MUSIGKRB06_1  Jkappa recombination signal bindin...   583  7.4e-75     1
• gp|M81867|MUSIGKRB03_1  Jkappa recombination signal bindin...   297  2.1e-34     1
• gp|M81875|MUSIGKRB10_1  Jkappa recombination signal bindin...   146  3.4e-21     2
• gp|M81873|MUSIGKRB08_1  Jkappa recombination signal bindin...   183  4.4e-18     1
  sp|P25172|SUZ2_DROME    SUPPRESSOR 2 OF ZESTE PROTEIN (PRO...   124  3.5e-13     2
  pir|S09358|S09358       notch protein - fruit fly (Drosoph...   134  9.5e-13     2
  sp|P36417|GBF_DICDI     G-BOX BINDING FACTOR. >pir|A53185|...   116  1.9e-12     2
  pir|A41726|A41726       homeotic protein BarH2 - fruit fly...    84  2.1e-11     4
  sp|P15330|DORS_DROME    EMBRYONIC POLARITY DORSAL PROTEIN....   122  2.5e-11     3
  gp|M13373|DROANTDFD_1   Dfd gene product [Drosophila melan...   107  2.9e-11     2
  pir|S16845|S16845       suppressor protein - fruit fly (Dr...   110  2.9e-11     2

500 lines omitted for brevity.

* gp|L12141|HUMFKHH_1     fork head-related protein [Homo sa...    73  0.28        1

Last 382 lines omitted for brevity.
```

Figure 2.5.13 Hit list from a BLASTP search of the fruit fly RBP-J_κ protein with no filtering. "Self hits" (marked by squares) correspond to other RBP-J_κ sequences present in the database. The sequence of the fork head–related protein is marked by a star.

Repeating the search with the same query and database, but with the addition of the filter seg + xnu option produces a much smaller hit list: 26 sequences containing one or more HSPs. The results are shown in Figure 2.5.14. The first eight hits are J_κ sequences (marked by squares in Figs. 2.5.13 and 2.5.14), followed by the fork head–related protein. Many of the false positives have disappeared as a result of the masking, allowing the fork head–related protein to improve its standing to ninth on the list. This example illustrates the necessity to use the seg + xnu filter systematically in protein similarity search.

LITERATURE CITED

Altschul, S.F. 1991. Amino acid substitution matrices from an information theoretic perspective. *J. Mol. Biol.* 219:555-565.

Altschul, S.F. 1993. A protein alignment scoring system sensitive at all evolutionary distances. *J. Mol. Biol. 36:290-300.*

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

Altschul, S.F., Boguski, M.S., Gish, W., and Wootton, J.C. 1994. Issues in searching molecular sequence databases. *Nature Genet.* 6:119-129.

Boguski, M.S., Lowe, T.M., and Tolstoshev, C.M. 1993. dbEST—Database for "expressed sequence tags." *Nature Genet.* 4:332-333.

Boguski, M.S., Tolstoshev, C.M., and Bassett, D.E. 1994. Gene discovery in dbEST. *Science* 265:1993-1994.

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Nordenskjold, M., Fishel, R., Kolodner, R., and Liskay, R.M. 1994. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* 368:258-261.

Brutlag, D.L. 1993. BLAZE: An implementation of the Smith-Waterman sequence comparison algorithm on a massively parallel computer. *Comput. Chem.* 17:203-207.

Claverie, J.-M. 1993. Detecting frameshifts by amino acid sequence comparison. *J. Mol. Biol.* 234:1140-1157.

Claverie, J.-M. and Makalowski, W. 1994. Letter. *Nature* 371:752.

Claverie, J.-M. and States, D.J. 1993. Information enhancement for large-scale sequence analysis. *Comput. Chem.* 17:191-201.

Doolittle, R.J. and Bork, P. 1993. Evolutionarily mobile modules in proteins. *Sci. Am.* 269:50-56.

Erickson, P., Gao, J., Chang, K.S., Look, T., Whisenant, E., Raimondi, S., Lasher, R., Trujillo, J., Rowley, J., and Drabkin, H. 1992. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation

BIU IAS / CENTRAL UNIVERSITY LIBRARY

```
                                                                  Smallest
                                                                    Sum
                                                          High   Probability
  Sequences producing High-scoring Segment Pairs:         Score    P(N)       N

• pir|A41585|A41585       J-kappa recombination signal-bindi...  2429  0.0        1
• sp|P28159|RBJK_DROME    J KAPPA-RECOMBINATION SIGNAL BINDI...  2360  0.0        2
• sp|P31266|RBJK_MOUSE    J KAPPA-RECOMBINATION SIGNAL BINDI...  1855  9.3e-256   1
• sp|Q06330|RBJK_HUMAN    J KAPPA-RECOMBINATION SIGNAL BINDI...  1838  2.3e-253   1
• gp|M81870|MUSIGKRB06_1 Jkappa recombination signal bindin...   583  6.1e-76    1
• gp|M81867|MUSIGKRB03_1 Jkappa recombination signal bindin...   258  2.9e-29    1
• gp|M81875|MUSIGKRB10_1 Jkappa recombination signal bindin...   146  1.5e-21    2
• gp|M81873|MUSIGKRB08_1 Jkappa recombination signal bindin...   183  2.5e-18    1
* gp|L12141|HUMFKHH_1     fork head-related protein [Homo sa...    73  0.22       1
  gp|M81876|MUSIGKRB11_1 Jkappa recombination signal bindin...    64  0.32       1
  sp|P13693|TCTP_HUMAN    TRANSLATIONALLY CONTROLLED TUMOR P...    50  0.60       2
  sp|P35584|HN3G_MOUSE    HEPATOCYTE NUCLEAR FACTOR 3-GAMMA ...    68  0.71       1
  sp|P32183|HN3G_RAT      HEPATOCYTE NUCLEAR FACTOR 3-GAMMA ...    68  0.71       1
  gp|Z25483|PESPKPRA_1    spike protein [Porcine epidemic di...    48  0.82       4
  gp|D26312|CHKESTFLTB_1 transrationally controlled tumor p...    50  0.82       2
  gp|L07874|HUMRBPJKC_1   recombination signal binding prote...    59  0.83       1
```

10 lines omitted for brevity.

```
>pir|A41585|A41585 J-kappa recombination signal-binding protein - fruit fly
           (Drosophila melanogaster) >pir|S15463|S15463 J-kappa recombination
           sequence-binding protein - fruit fly (Drosophila melanogaster)
           >gp|X58393|DMDRBP_1 Jkappa-recombination signal binding protein
           [Drosophila melanogaster]
           Length = 550

 Score = 2429 (1128.2 bits), Expect = 0.0, P = 0.0
 Identities = 466/550 (84%), Positives = 466/550 (84%)

Query:      1 MPHFGLPGPQPPSSXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXMSLLPGPYRPHIEEKK 60
              MPHFGLPGPQPPSS                               MSLLPGPYRPHIEEKK
Sbjct:      1 MPHFGLPGPQPPSSQQQQQQLQVHHQQQQQQQQQQQQQQHQQQMQMSLLPGPYRPHIEEKK 60

Query:     61 LTRDAMEKYMRDGNDMVIVILHAKVAQKSYGNEKRFFCPPPCIYLFGSGWRRRYXXXXXX 120
              LTRDAMEKYMRDGNDMVIVILHAKVAQKSYGNEKRFFCPPPCIYLFGSGWRRRY
Sbjct:     61 LTRDAMEKYMRDGNDMVIVILHAKVAQKSYGNEKRFFCPPPCIYLFGSGWRRRYEEMLQQ 120
```

Alignment from residues 121 to 540 omitted for brevity.

```
Query:    541 QWNSHGSGLS 550
              QWNSHGSGLS
Sbjct:    541 QWNSHGSGLS 550
```

7 alignments omitted for brevity.

```
>gp|L12141|HUMFKHH_1 fork head-related protein [Homo sapiens]
           Length = 347

 Score = 73 (33.9 bits), Expect = 0.24, P = 0.22
 Identities = 15/28 (53%), Positives = 18/28 (64%)

Query:    364 DKAEYQFYEGMGPVASPVTPVPIVNSLN 391
              D AE+ +Y   G V SPVTPVP +  LN
Sbjct:     11 DLAEWSYYPEAGEVYSPVTPVPTMAPLN 38
```

Remaining alignments omitted for brevity.

Figure 2.5.14 Hit list for a BLASTP search of fruit fly J_K protein using seg + xnu filtering. Several "self hits" (marked by squares) correspond to other RBP-J_K sequences present in the database. The sequence of the *H. sapiens* fork head–related protein (marked by a star) was ranked 516th in the initial search and improved its standing to ninth with filtering.

of a fusion transcript, AML1/ETO, with similarity to *Drosophila* segmentation gene, runt. *Blood* 80:1825-1831.

Feigl, G., Gram, M., and Pongs, O. 1989. A member of the steroid hormone receptor gene family is expressed in the 2O-OH-ecdysone inducible puff 75B in *Drosophila melanogaster*. *Nucl. Acids Res.* 17:7167-7178.

Furukawa, T., Kawaichi, M., Matsunami, N., Ryo, H., Nishida, Y., and Honjo, T. 1991. The *Drosophila* RBP-J kappa gene encodes the binding protein for the immunoglobulin J kappa recombination signal sequence. *J. Biol. Chem.* 266:23334-23340.

Gish, W. and States, D.J. 1993. Identification of protein coding regions by database similarity search. *Nature Genet.* 3:266-272.

BOSTON CENTRAL UNIVERSITY LIBRARY

Henikoff, S. and Henikoff, J.G. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U.S.A.* 89:10915-10919.

Hoey, T., Weinzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D., and Tjian, R. 1993. Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* 72:247-260.

Hromas, R., Moore, J., Johnston, T., and Socha Klemsz, M. 1993. *Drosophila* forkhead homologues are expressed in a lineage-restricted manner in human hematopoietic cells. *Blood* 81:2854-2859.

Jurka, J., Walichiewicz, J., and Milosavljevic, A. 1992. Prototypic sequences for human repetitive DNA. *J. Mol. Evol.* 35:286-291.

Karlin, S. and Altschul, S.F. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. U.S.A.* 87:2264-2268.

Karlin, S. and Altschul, S.F. 1993. Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci. U.S.A.* 90:5873-5877.

Kokubo, T., Gong, D.W., Roeder, R.G., Horikoshi, M., and Nakatani, Y. 1993. The *Drosophila* 110-kDa transcription factor TFIID subunit directly interacts with the N-terminal region of the 230-kDa subunit. *Proc. Natl. Acad. Sci. U.S.A.* 90:5896-5900.

McKusick, V. 1994. Online Mendelian Inheritance in Man. Johns Hopkins University Press, Baltimore.

Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Masekim, N., Kanekom, Y., Kamada, N., and Ohki, M. 1993. The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. *EMBO J.* 12:2715-2721.

Owens, G.P., Hahn, W.E., and Cohen, J.J. 1991. Identification of mRNAs associated with programmed cell death in immature thymocytes. *Mol. Cell Biol.* 11:4177-4188.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., Venter, J.C., Hamilton, S.R., Peterson, G.M., Watson, P., Lynch, H.T., Peltomaki, P., Mecklin, J.-P., de la Chapelle, A., Kinzler, K.W., and Vogelstein, B. 1994. Mutation of a *mutL* homolog in hereditary colon cancer. *Science* 263:1625-1629.

Patthy, L. 1991. Modular exchange principles in proteins. *Curr. Opin. Struct. Biol.* 1:351-361.

Pearson, W.R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63-98.

Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85:2444-2448.

Smith, M.W., Holmsen, A.L., Wei, Y.H., Peterson, M., and Evans, G.A. 1994. Genomic sequence sampling: A strategy to high-resolution sequence-based physical mapping of complex genomes. *Nature Genet.* 7:40-47.

Smith, T.F. and Waterman, M.S. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* 147:195-197.

States, D.J., Gish, W., and Altschul, S.F. 1991. Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. *Methods* 3:66-70.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Smaldon, N., Smith, A., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. 1994. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* 368:32-38.

Wootton, J.C. 1994a. Non-globular domains in protein sequences: Automated segmentation using complexity measures. *Comput. Chem.* 18:269-285.

Wootton, J.C. 1994b. Simple sequences of protein and DNA. *In* Nucleic Acid and Protein Sequence Analysis: A Practical Approach, 2nd ed. (M.J. Bishop and C.R. Rawlings, eds.) IRL Press, Oxford.

Wootton, J.C. and Federhen, S. 1993. Statistics of low complexity in amino acid sequences and sequence databases. *Comput. Chem.* 17:149-163.

Contributed by Hervé E. Recipon, Gregory D. Schuler, and Mark S. Boguski
National Center for Biotechnology Information
Bethesda, Maryland

Warren Gish
Washington University School of Medicine
St. Louis, Missouri

BIU IAS CENTRAL UNIVERSITY LIBRARY

DETECTION AND ASSAY METHODS

3

INTRODUCTION 3.0.1

3.1 Spectrophotometric Determination of Protein Concentration 3.1.1
Basic Protocol 1: Using A_{280} to Determine Protein Concentration 3.1.1
Alternate Protocol: Using A_{205} to Determine Protein Concentration 3.1.2
Basic Protocol 2: Using Fluorescence Emission to Determine Protein Concentration 3.1.3
Reagents and Solutions 3.1.3
Commentary 3.1.4

3.2 Quantitative Amino Acid Analysis 3.2.1
Information Obtained from Amino Acid Analysis 3.2.1
Sample Preparation 3.2.2
Calculation of the Average Composition 3.2.2

3.3 In Vitro Radiolabeling of Peptides and Proteins 3.3.1
Basic Protocol 1: Iodination at Tyrosine or Histidine Residues Using Iodo-Beads 3.3.3
Alternate Protocol 1: Iodination at Tyrosine or Histidine Residues Using Chloramine T or Iodogen 3.3.5
Alternate Protocol 2: Iodination at Tyrosine or Histidine Residues Using Lactoperoxidase 3.3.6
Support Protocol 1: Equimolar Iodination to Give High Yields of Product 3.3.7
Support Protocol 2: Separation of Unlabeled Peptide from Iodinated Products by HPLC 3.3.8
Basic Protocol 2: Iodination at Lysine Residues or N-Terminus Using Bolton-Hunter Reagent 3.3.9
Basic Protocol 3: ^{14}C or ^{3}H Labeling at Lysine Residues or N-Terminus by Acetylation Using Anhydride 3.3.11
Alternate Protocol 3: ^{14}C or ^{3}H Labeling at Lysine Residues or N-Terminus by Reductive Alkylation 3.3.12
Alternate Protocol 4: ^{14}C or ^{3}H Labeling at Cysteine Residues Using Iodoacetic Acid or Iodoacetamide 3.3.13
Basic Protocol 4: ^{3}H Labeling at Tyrosine Residues by Reduction of ^{127}I-Labeled Peptide 3.3.13
Basic Protocol 5: Introduction of Labeled Amino Acid Residues During Peptide Synthesis 3.3.14
Alternate Protocol 5: Selective Labeling of N-Terminus During Peptide Synthesis 3.3.15
Commentary 3.3.16

3.4 Assays for Total Protein 3.4.1
Basic Protocol 1: Biuret Assay for Quantitation of Total Protein 3.4.1
Basic Protocol 2: Hartree-Lowry Assay for Quantitation of Total Protein 3.4.3
Basic Protocol 3: Bicinchinoic Acid (BCA) Assay for Quantitation of Total Protein 3.4.5
Basic Protocol 4: Acid Digestion–Ninhydrin Method for Quantitation of Total Protein 3.4.7
Support Protocol 1: Heat Sealing Glass Tubes 3.4.8

continued

BCU IAN / CENTRAL UNIVERSITY LIBRARY

Copyright © 1996 by John Wiley & Sons, Inc.

Basic Protocol 5: Ultraviolet Absorption to Measure Total Protein 3.4.9
Basic Protocol 6: Coomassie Dye–Binding Assay (Bradford Assay) to Measure Total Protein 3.4.10
Basic Protocol 7: Dry Weight Determination to Measure Total Protein 3.4.11
Support Protocol 2: Gel Dialysis of Protein Samples 3.4.12
Support Protocol 3: Trichloroacetic Acid Precipitation of Protein Samples 3.4.13
Reagents and Solutions 3.4.13
Commentary 3.4.15

3.5 Kinetic Assay Methods
General Aspects of Assay Design 3.5.1
Continuous Versus Stop-Time Assays 3.5.1
Coupled Enzyme Assays 3.5.7
Binding Studies 3.5.8
Presentation of Initial-Rate Data 3.5.8
Concluding Remarks 3.5.9
Literature Cited 3.5.9
Key References 3.5.10

3.6 Biotinylation of Proteins in Solution and on Cell Surfaces 3.6.1
Basic Protocol 1: Covalent Attachment of Biotin to Lysines 3.6.1
Basic Protocol 2: Side-Specific Biotinylation of Proteins on Cell Surfaces 3.6.4
Basic Protocol 3: Covalent Attachment of Biotin to Thiols 3.6.6
Support Protocol 1: Reduction of Disulfide Bonds 3.6.7
Basic Protocol 4: Covalent Attachment of Biotin to Carbohydrates 3.6.8
Support Protocol 2: Detection of Biotinylated Proteins 3.6.10
Reagents and Solutions 3.6.10
Commentary 3.6.11

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 3
Detection and Assay Methods

INTRODUCTION

The success or failure of protein-centered projects can frequently be traced to the quality of the analytical procedures used to characterize the sample at different stages. Qualitative and quantitative analysis can aid in definition of the sample for the purpose of designing separations. A knowledge of the properties of the desired protein (e.g., whether it has a high aromatic amino acid content) can suggest methods of analysis that will help locate the desired protein in a complex mixture. Establishing the properties of an isolated protein creates benchmarks that future researchers can use to evaluate their protocols and final product. Accurate quantitation of the amount of protein at the beginning, middle, and end of a series of steps is the only valid way to evaluate the overall yield of a procedure. The observation of significant loss of protein following a particular purification procedure, without a substantial increase in the purity of the desired protein, would indicate that the procedure should be omitted or revised.

Several spectroscopic procedures for characterizing protein samples are described in *UNIT 3.1*. Measuring the absorbance of the aromatic amino acids in a protein at different wavelengths yields a very useful measure of protein concentration. This is nondestructive and requires very little sample or time. A more qualitative, but much more sensitive, evaluation is provided by fluorescence spectroscopy.

Another approach is amino acid analysis—qualitative analysis, to determine purity, and quantitative analysis, to provide concentrations—both of which are presented in *UNIT 3.2*. Procedures are also given for calculating amino acid ratios from primary analytical data. Significant advances that have improved the precision and sensitivity of amino acid analysis have reinvigorated this method, which had for some years been neglected.

A common method for enhancing the sensitivity of protein detection is to radiolabel the proteins, as detailed in *UNIT 3.3*. This approach may also be employed to help quantitate the binding of peptide to other molecules or for radioimmunoassays. Several procedures for iodination (introduction of ^{125}I or ^{131}I) are presented; the choice between them will depend on the sample type, the content of residues sensitive to oxidation, and the amino acids that are critical to function. Procedures for handling radioactive samples, for separating labeled from unlabeled peptides, and for iodination to equimolar or higher levels (using a mixture of radioactive and nonradioactive iodine isotopes) are also provided. When Tyr or His residues are absent, or cannot be iodinated because they are critical for function, alternative methods involving the incorporation of ^{14}C or ^{3}H may be employed instead. The methods described include labeling by acetylation at amino groups on lysine or at the amino terminus using the Bolton-Hunter reagent or acetic anhydride, by reductive alkylation, and by labeling at Lys residues with iodoacetic acid or iodoacetamide. Two synthetic methods, reduction of a peptide containing iodo-Tyr residues and incorporation of labeled amino acids, are also presented.

Measurement of the total protein content of a sample is discussed in *UNIT 3.4*. Three methods based on reduction of Cu^{2+} to Cu^{+} followed by chelation of the cuprous ion by various reagents (biuret, Hartree-Lowry, and bicinchinonic acid) are presented. In addition, a quantitative measure of total protein based on protein hydrolysis and the ninhydrin

Contributed by Ben M. Dunn
Current Protocols in Protein Science (1997) 3.0.1-3.0.2
Copyright © 1997 by John Wiley & Sons, Inc.

BOWMAN GENERAL UNIVERSITY LIBRARY

reaction is described. The use of UV absorption, Coomassie dye binding, and weight analysis to determine protein content complete the presentation. Important considerations for performing these assays, such as dealing with interfering substances, are detailed throughout the unit.

Some aspects involved in the design of enzyme assays are discussed in *UNIT 3.5*. Parameters that affect both enzyme and substrate are evaluated and suggestions to optimize the quantitation of enzymatic activity are enumerated. The proper concentration range for conducting initial velocity measurements is presented, and different methods of data presentation and analysis are discussed.

Tagging a protein with a group that permits facile detection is an excellent way, for example, to study the movement of a protein through internalization or compartmentalization. A variety of reagents are available for attaching biotin to functional groups or the surface of a protein (*UNIT 3.6*). Detection and localization can be achieved by the strong and specific binding of avidin or streptavidin. These reactions can also be used to label proteins at cell surfaces.

Additional procedures to quantitate total protein content by biological activity by analytical assay methods will be presented in future updates to this chapter.

Ben M. Dunn

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Spectrophotometric Determination of Protein Concentration

This unit describes spectrophotometric methods for measuring the concentration of a sample protein in solution. In Basic Protocol 1, absorbance measured at 280 nm (A_{280}) is used to calculate protein concentration by comparison with a standard curve or published absorptivity values for that protein (a_{280}). In the Alternate Protocol, absorbance measured at 205 nm (A_{205}) is used to calculate the protein concentration. The A_{280} and A_{205} methods can be used to quantitate total protein in crude lysates and purified or partially purified protein. Both of these methods are simple and can be completed quickly. The A_{280} method is the most commonly used. The A_{205} method can detect lower concentrations of protein and is useful for dilute protein samples, but is more susceptible to interference by reagents in the protein sample than the A_{280} method. Basic Protocol 2 uses a spectrofluorometer or a filter fluorometer to measure the intrinsic fluorescence emission of a sample solution; this value is compared with the emissions from standard solutions to determine the sample concentration. The fluorescence emission method is used to quantitate purified protein. This simple method is useful for dilute protein samples and can be completed in a short amount of time.

USING A_{280} TO DETERMINE PROTEIN CONCENTRATION

BASIC PROTOCOL 1

Determination of protein concentration by measuring absorbance at 280 nm (A_{280}) is based on the absorbance of UV light by aromatic amino acids in protein solutions—due primarily to tryptophan and tyrosine residues and to a lesser extent phenylalanine residues. The measured absorbance of a protein sample solution is used to calculate the concentration either from its published absorptivity at 280 nm (a_{280}) or by comparison with a calibration curve prepared from measurements with standard protein solutions. This assay can be used to quantitate solutions with protein concentrations of 20 to 3000 µg/ml.

Materials

3 mg/ml standard protein solution (see recipe; optional)
Sample protein
Spectrophotometer with UV lamp

1. For calibrating with standards, use the 3 mg/ml standard protein solution to prepare dilutions of 20, 50, 100, 250, 500, 1000, 2000, and 3000 µg/ml in the same solvent as used to prepare the sample protein. Prepare a blank consisting of solvent alone.

   Ideally, for purified or partially purified protein, the protein standard should have an aromatic amino acid content similar to that of the sample protein. For the total protein of a crude lysate, bovine serum albumin (BSA) is a commonly used standard for spectrophotometric quantitation of protein concentration. A 3 mg/ml solution of BSA should have an A_{280} of 1.98, based on an A_{280} of 6.61 for a 1% (w/v) solution.

2. Turn on the UV lamp of the spectrophotometer and set the wavelength to 280 nm. Allow the instrument to warm up 30 min before taking measurements.

3. Zero the spectrophotometer with the solvent blank.

4. Measure the absorbance of the protein standard and unknown solutions.

   If the A_{280} of the sample protein is >2.0, dilute the sample further in the same solvent and measure the A_{280} again.

Copyright © 1996 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

5a. *If the a_{280} of the protein is known:* Calculate the unknown sample concentration from its absorbance value using Equation 3.1.1.

In this equation a_{280} has units of ml/mg cm and b is the path length in cm.

$$\text{concentration (mg/ml)} = \frac{A_{280}}{a_{280} \times b}$$

Equation 3.1.1

5b. *If standard solutions are used for quantitation:* Create a calibration curve by either plotting or performing regression analysis of the A_{280} versus concentration of the standards. Use the absorbance of the sample protein to determine the concentration from the calibration curve.

ALTERNATE PROTOCOL

USING A_{205} TO DETERMINE PROTEIN CONCENTRATION

Determination of protein concentration by measurement of absorbance at 205 nm (A_{205}) is based on absorbance by the peptide bond. The concentration of a protein sample is determined from the measured absorbance and the absorptivity at 205 nm (a_{205}). This assay can be used to quantitate protein solutions with concentrations of 1 to 100 µg/ml protein.

Additional Materials (also see Basic Protocol 1)

Brij 35 solution: 0.01% (v/v) Brij 35 (Sigma) in an aqueous solution appropriate for dissolving or diluting the sample protein

1. Dissolve or dilute the protein sample in Brij 35 solution.
2. Turn on the UV lamp of the spectrophotometer and set the wavelength to 205 nm. Allow the instrument to warm up 30 min before taking measurements.
3. Zero the spectrophotometer with the Brij 35 solution alone.
4. Measure the absorbance of the sample protein.

5a. *If the a_{205} of the protein is known:* Use Equation 3.1.1 to calculate the concentration of the sample protein *except* substitute the appropriate values for A_{205} and a_{205}.

5b. *If the a_{205} is not known:* Estimate the concentration of the sample protein from its measured absorbance using Equation 3.1.2.

$$\text{concentration (mg/ml)} = \frac{A_{205}}{31 \times b}$$

Equation 3.1.2

In this equation, the absorptivity value, 31, has units of ml/mg cm and b is the path length in cm.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

USING FLUORESCENCE EMISSION TO DETERMINE PROTEIN CONCENTRATION

Protein concentration can also be determined by measuring the intrinsic fluorescence based on fluorescence emission by the aromatic amino acids tryptophan, tyrosine, and/or phenylalanine. Usually tryptophan fluorescence is measured. The fluorescence intensity of the protein sample solution is measured and the concentration calculated from a calibration curve based on the fluorescence emission of standard solutions prepared from the purified protein. This assay can be used to quantitate protein solutions with concentrations of 5 to 50 μg/ml.

Materials

Protein standard solution prepared using the purified protein (see recipe)
Sample protein
Spectrofluorometer *or* filter fluorometer with an excitation cutoff filter ≤285 nm and an emission filter >320 nm

1. Prepare dilutions of the purified protein at 5, 7.5, 10, 25, and 50 μg/ml in the same solvent as the sample protein. Prepare a blank consisting of solvent alone.
2. Turn on the lamp of the instrument and allow it to warm up 30 min before taking measurements.

   If a spectrofluorometer is used, set the excitation wavelength to 280 nm and the emission wavelength to between 320 and 350 nm. If the exact emission wavelength is not known, determine it empirically by scanning the standard solution with the excitation wavelength set to 280 nm. If the instrument is a filter fluorometer, use an excitation cutoff filter ≤285 nm and an emission filter >320 nm.
3. Zero the instrument with the solvent blank.
4. Measure the fluorescence of the protein standard and sample protein solutions.
5. Create a calibration curve by either plotting or performing regression analysis of the fluorescence intensity versus concentration of the standards. Using the fluorescence intensity of the sample protein, determine the concentration from the calibration curve.

   Fluorescence emission is a linear function of concentration only over a limited range.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Standard protein solution, 3 mg/ml

Weigh out dry protein and prepare a stock solution at a concentration of 3 mg/ml in the same solvent as used for the sample protein. Store up to 3 months at –20°C.

To prepare calibration standard solutions, dilute the stock solution in solvent to give the desired final concentrations for the standard curve.

Bovine serum albumin (BSA, fraction V; Sigma) is frequently used for a protein standard solution. A 3 mg/ml solution of BSA should have an A_{280} of 1.98, based on an A_{280} of 6.61 for a 1% (w/v) solution.

For quantitation of a purified or partially purified protein, if possible, the protein standard should have an aromatic amino acid content similar to that of the sample protein.

BOULDER CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

Measuring absorbance at 280 nm (A_{280}) is one of the oldest methods for determining protein concentration (Warburg and Christian, 1942; Layne, 1957). This method is still widely used because it is simple and does not require incubating the sample with exogenous chromophores. However, the detection limit is higher than colorimetric methods and therefore higher concentrations of protein are necessary. The A_{280} method requires that the protein being quantitated have aromatic amino acids, primarily tryptophan. Because of the variability in aromatic amino acid content among different proteins, their absorptivity at 280 nm (a_{280}) also varies. Therefore, if calibration standards are used for quantitation, the aromatic amino acid content of the standard must be similar to that of the sample protein for accurate results. Accordingly, the quantitation of proteins by peptide bond absorption at 205 nm (A_{205}) is more universally applicable among proteins. Furthermore, the absorptivity for a given protein at 205 nm is several-fold greater than that at 280 nm (Scopes, 1974; Stoscheck, 1990). Thus lower concentrations of protein can be quantitated with the A_{205} method. The disadvantage of this method is that some buffers and other components absorb at 205 nm (Stoscheck, 1990).

In addition to the aromatic amino acids, several others have absorption maxima in the UV range. Table 3.1.1 shows the wavelengths of absorption maxima and corresponding molar absorptivity (ε) for the amino acids with appreciable absorbance in the UV range. Only tryptophan has an absorption maximum at 280 nm, although tyrosine will absorb somewhat at that wavelength. The ε_{280} for tryptophan is nearly five-fold greater than that for tyrosine (Table 3.1.2). Several amino acids other than those in Table 3.1.1 absorb light below 205 nm (Fasman, 1989), but either the molar absorptivities are too low to be significant or the wavelengths are too short for practical absorbance measurements.

The aromatic amino acids also have fluorescence emissions when excited by light in the UV range. Table 3.1.3 gives the excitation wavelength, fluorescence emission wavelength, and quantum yield (Q) for tryptophan, tyrosine, and phenylalanine. The quantum yield is the ratio of photons emitted to photons absorbed. Typically, phenylalanine fluorescence is not detected in the presence of tyrosine and tryptophan due to low Q. Furthermore, tyrosine fluorescence is nearly completely quenched if the tyrosine residue is ionized or near an amino group, a carboxyl group, or a tryptophan residue (Teale, 1960; Freifelder, 1982). Therefore, tryptophan fluorescence is what is customarily measured.

Measurement of intrinsic fluorescence by aromatic amino acids is primarily used to obtain qualitative information (Freifelder, 1982). However, with a protein standard whose aromatic amino acid content is similar to that of the sample, intrinsic fluorescence can be used for quantitation (Hawkins and Honigs, 1987). An additional consideration is that the tertiary

Table 3.1.1 Absorption Maxima and Molar Absorptivity (ε) of Amino Acids[a]

Amino acid	Wavelength maxima (nm)	$\varepsilon \times 10^{-3}$ (l/mol cm)
Cysteine	250	0.3
Histidine	211	5.9
Phenylalanine	188	60.0
	206	9.3
	257	0.2
Tryptophan	219	47.0
	279	5.6
Tyrosine	193	48.0
	222	8.0
	275	1.4

[a]Values are for aqueous solutions at pH 7.1 (Freifelder, 1982; Fasman, 1989).

RCU IASI // CENTRUL UNIVERSITAR

Table 3.1.2 Molar Absorptivity (ε) of Aromatic Amino Acids at 280 nm[a]

Amino acid	$\varepsilon \times 10^{-3}$ (l/mol cm)
Phenylalanine	0.0007
Tryptophan	5.559
Tyrosine	1.197

[a]Values are for aqueous solutions at pH 7.1 (Fasman, 1989).

Table 3.1.3 Fluorescence Properties of Aromatic Amino Acids[a]

Amino acid	Excitation wavelength	Emission wavelength	Quantum yield
Phenylalanine	260 nm	283 nm	0.04
Tryptophan	285 nm	360 nm	0.20
Tyrosine	275 nm	310 nm	0.21

[a]Values are for aqueous solutions at pH 7 and 25°C (Hawkins and Honigs, 1987; Fasman, 1989).

structure of a protein will influence the fluorescence, e.g., adjacent protonated acidic groups in a protein molecule will quench tryptophan fluorescence (Freifelder, 1982).

Critical Parameters and Troubleshooting

A 1-cm path length quartz cuvette is most often used to make absorbance measurements. However, quartz cuvettes with shorter path lengths, 0.01 to 0.5 cm, are available (e.g., from Hellma Cells or Beckman); these shorter path length cuvettes allow higher concentrations of protein solutions to be measured. Equation 3.1.1 and Equation 3.1.2 assume the cuvette has a path length of 1 cm; when cuvettes of shorter path length are used, the correct value for b must be substituted in the equation.

The solvent pH and polarity will affect the absorbance and fluorescence properties of a protein. A notable example of pH effects on absorbance is seen with tyrosine residues, where a change in pH from neutral to alkaline results in a shift of the absorbance maximum to a longer wavelength and an increase in absorptivity due to dissociation of the tyrosine phenolic hydroxyl group (Freifelder, 1982; Fasman, 1989). An example of solvent polarity effects on fluorescence is observed with tryptophan, where a decrease in solvent polarity results in a shift in fluorescence emission to shorter wavelengths and an increase in intensity (Freifelder, 1982). Because of these effects, the following precautions should be taken for accurate results: (1) when calibration curves are used for quantitation by absorbance or fluorescence, standards must be in the same solvent as the samples; and (2) when a published absorptivity at a given wavelength is used for quantitation, the solvent composition of the sample must be the same as that used in obtaining the published data.

Many buffers and other reagents can interfere with A_{280} and A_{205} spectrophotometric measurements. Stoscheck (1990) lists the concentration limits for many such reagents used in these spectrophotometric methods. The more commonly used reagents that absorb at 280 and 205 nm are listed in Table 3.1.4. In addition, reagents that contain carbon-carbon or carbon-oxygen double bonds can interfere with the A_{205} method.

Because stray light can affect the linearity of absorbance versus concentration, absorbance values >2.0 should not be used for sample proteins measured by the A_{280} or A_{205} method. Samples with absorbance >2.0 should be diluted further in the appropriate buffer to obtain absorbances <2.0.

Nucleic acids have substantial absorbance at 280 nm and can interfere with A_{280} quantitation of protein in crude samples. To resolve the protein concentration in such samples, measure the absorbance at 260 nm and 280 nm and calculate the protein concentration as follows (Warburg and Christian, 1942; Layne, 1957): protein concentration (mg/ml) = $1.55 \times A_{280} - 0.76 \times A_{260}$.

BOSTON UNIVERSITY CENTRAL LIBRARY

Table 3.1.4 Concentration Limits of Interfering Reagents for A_{205} and A_{280} Protein Assays[a]

Reagents[b]	A_{205}	A_{280}
Ammonium sulfate	9% (w/v)	>50% (w/v)
Brij 35	1% (v/v)	1% (v/v)
DTT	0.1 mM	3 mM
EDTA	0.2 mM	30 mM
Glycerol	5% (v/v)	40% (v/v)
KCl	50 mM	100 mM
2-ME	<10 mM	10 mM
NaCl	0.6 M	>1 M
NaOH	25 mM	>1 M
Phosphate buffer	50 mM	1 M
SDS	0.10% (w/v)	0.10% (w/v)
Sucrose	0.5 M	2 M
Tris buffer	40 mM	0.5 M
Triton X-100	<0.01% (v/v)	0.02% (v/v)
TCA	<1% (w/v)	10% (w/v)
Urea	<0.1 M	>1 M

[a]Values from Stoscheck, (1990).

[b]Abbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; 2-ME, 2-mercaptoethanol; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid.

This estimation of protein concentration is valid up to 20% (w/v) nucleic acid or an A_{280}/A_{260} ratio <0.6.

Anticipated Results

Depending on the protein, the concentration range for the A_{280} method is 20 to 3000 µg/ml, for the A_{205} method is 1 to 100 µg/ml, and for the fluorescence emission method is 5 to 50 µg/ml.

Published absorptivities of proteins at 280 nm are usually given as the absorbance for a 1% (w/v) protein solution per cm, $A^{1\%}$, or as the molar absorptivity, ε, which has units of l/mol cm. To convert these published coefficients to units of mg/ml, use either Equation 3.1.3 or Equation 3.1.4.

$$\text{concentration (mg/ml)} = \frac{A_{280} \times 10}{A^{1\%} \times b}$$

Equation 3.1.3

Molecular weight in Equation 3.1.4 is the molecular weight of the protein.

$$\text{concentration (mg/ml)} = \frac{A_{280} \times \text{molecular weight}}{\varepsilon_{280} \times b}$$

Equation 3.1.4

Time Considerations

When the absorptivity for a protein is known, the A_{280} and A_{205} measurements require <30 min depending on the number of samples. When standards are used for quantitation with these assays or for intrinsic fluorescence quantitation, 1 hr is required.

Literature Cited

Fasman, G.D. 1989. Practical Handbook of Biochemistry and Molecular Biology. CRC Press, Boca Raton, Fla..

Freifelder, D. 1982. Physical Biochemistry: Applications to Biochemistry and Molecular Biology 2nd ed. W.H. Freeman, New York.

Hawkins, B.K. and Honigs, D.E. 1987. A comparison of spectroscopic techniques for protein quantification in aqueous solutions. *Am. Biotechnol. Lab.* 5:26-37.

ECU IASTATE CENTRAL UNIVERSITY LIBRARY

Layne, E. 1957. Spectrophotometric and turbidimetric methods for measuring proteins. *Methods Enzymol.* 3:447-454.

Scopes, R.K. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* 59:277-282.

Stoscheck, C.M. 1990. Quantitation of protein. *Methods Enzymol.* 182:50-68.

Teale, F.W.J. 1960. The ultraviolet fluorescence of proteins in neutral solutions. *Biochem. J.* 76:381-388.

Warburg, O. and Christian, W. 1942. Isolierung und kristallisation des garungsferments enolase. *Biochem. Z.* 310:384-421.

Key References

Chen, R.F. 1990. Fluorescence of proteins and peptides. *In* Practical Fluorescence, 2nd ed. (G.G. Guilbault, ed.) pp. 575-682. Marcel Dekker, Inc., New York.

Detailed discussion of intrinsic fluorescence of proteins and what factors affect fluorescence emission by the aromatic amino acids (see pp. 618-663).

Fasman, G.D. 1989. See above.

Contains tables with absorptivities for UV spectrophotometric detection and tables with data on excitation and emission wavelengths for fluorescence detection of many proteins. Also includes a table with molecular weights for many characterized proteins.

Stoscheck, C.M. 1990. See above.

Contains list of substances that can interfere with 205- and 280-nm spectrophotometric measurements of proteins and of concentration limits for these substances.

Contributed by Michael H. Simonian
Beckman Instruments, Inc.
Fullerton, California

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Quantitative Amino Acid Analysis

The use of quantitative amino acid analysis in the characterization of protein samples is often overlooked in many laboratories. This was typically due to the difficulties involved in obtaining reliable amino acid composition data, as well as to the advent of colorimetric dye-binding assays as alternative methods to determine protein concentration. However, many proteins do not bind dyes in the same way as the proteins normally used as standards, making quantitation by those techniques difficult. Fortunately, at many institutions, the creation of protein chemistry core laboratories devoted mainly to protein sequence analysis and peptide synthesis has provided new instrumentation that yields the necessary precision. This unit describes the information that can be derived from quantitative amino acid analysis, presents some details on sample preparation, and gives examples of the calculations that are needed.

INFORMATION OBTAINED FROM AMINO ACID ANALYSIS

The benefits of amino acid composition data are two-fold. First, identifying the nature of the amino acids present in a sample can provide an important qualitative measure of the purity of the preparation. Amino acid analysis does not depend on the shape, charge, or function of a protein because the hydrolysis completely destroys structure and reduces the protein to its constituent amino acids. Although large proteins typically contain a mixture of all the commonly occurring amino acids, certain residues are relatively rare—Trp, Met, His, Cys, and occasionally Arg are present at lower than average ratios compared to Gly, Ala, Ser, and Leu. For a pure protein, the ratios of the amino acids should approach the theoretical levels to within ~10%, with a few caveats. For a sample from a column fraction, for example, the presence of a higher than expected ratio of a certain amino acid is a very good indication that the sample is not pure and that further fractionation is indicated. In the same way, a lower than expected ratio could also indicate impurity—for example, the presence of a second protein that has a lower ratio of some significant amino acid present in higher amounts in the protein of interest. Because the expression of a protein in bacteria or other hosts utilizes a gene of known sequence, the expected amino acid composition is also known (see Chapters 5 and 6 for expression and purification of recombinant proteins).

Second, averaged composition yields an excellent measure of the amount of protein in a sample. On the ion-exchange-based Beckman 6300 amino acid analyzer running with ninhydrin detection chemistry, the minimum amount of the standard (a mixture of amino acids used to derive integration parameters as well as elution times) that is required is typically 1 nmol of each amino acid. Amounts as low as 250 pmol (0.25 nmol) of an amino acid can be measured with confidence from protein hydrolyzates. For a protein of 50 kDa mass, 1 nmol is 50 μg, an amount that can be obtained readily by purification of the protein from a high-expression system followed by purification by standard methods. For isolation of a trace protein from a natural source, this required amount may be too large to be feasible. However, a protein of 50 kDa mass contains ~400 amino acids, or roughly 20 copies of each amino acid on average. Thus, the actual amount of protein required for detection and semi-quantitative analysis could be as low as 0.05 nmol, or 2.5 μg. At that level, rare amino acids cannot be accurately quantitated, but the more common ones can be measured easily. Obviously, the more material available, the more accurate the analysis. Users should check with their local facility for specific advice on the amount of sample that can be processed by existing technologies.

The advent of advanced instrumentation, such as the Applied Biosystems 420 amino acid analyzer and the Waters PicoTag system, has provided another increment in sensitivity. These instruments utilize precolumn derivatization followed by separation based on reversed-phase high-performance liquid chromatography (HPLC). With such analyzers, 100 pmol of each amino acid, or 5 pmol of the 50-kDa protein described above, can be quantitated accurately.

Another important consideration in the design of a purification scheme that can be accomplished through the use of amino acid quantitation is the compilation of the mass balance—the relationship between how much protein is applied to the chromatography column and how much is recovered in a particular step. Amino acid analysis of a measured aliquot of the sample taken before application (typically ≤1%) along with aliquots of column frac-

Copyright © 1995 by John Wiley & Sons, Inc.

BOWLING GREEN STATE UNIVERSITY LIBRARY

tions or pools, can be used to calculate the amounts of protein loaded and recovered. These calculations can provide essential information to assess the success of a particular purification procedure.

SAMPLE PREPARATION

The success or failure of quantitative amino acid analysis is strongly dependent upon the quality of the sample, as well as upon the exact composition of the sample matrix (i.e., the aggregate of all components including salts, detergent, and dirt). The following section gives a few key recommendations for successful sample preparation.

Sample Concentration

An ideal sample for amino acid analysis would contain 1 to 10 nmol protein in a volume of 10 to 100 µl water. A sample that has been dialyzed against a low concentration of buffer (if needed to maintain solubility) or pure water is optimal, because the presence of salt, detergent, high concentrations of buffers, or other agents may interfere with either hydrolysis or subsequent analysis.

A sample that has been lyophilized to dryness is also usually acceptable, with some important considerations. First, lyophilization should be done in such a way as to concentrate the lyophilized residue at the bottom of a conical tube. Drying a sample in a round-bottom flask may leave the protein sample spread over a large surface area. Such samples are extremely difficult to handle and will not give satisfactory results. Second, lyophilization of a buffered solution will concentrate the salts as well as the protein. Therefore, it is recommended that the total ionic strength of the solution be <50 µM before lyophilization.

Estimate of Total Protein

An estimate of the total amount of protein present is very helpful in planning an accurate analysis. This can be based on a colorimetric dye–binding assay, an A_{280} measurement, or an activity assay in which the results are compared to a standard value.

All information about the sample should be communicated to the individual charged with conducting the analysis to ensure the maximum return on the sample. Many facilities provide a form with a series of questions to be answered concerning the sample.

Samples from PVDF Blots

With the advent of more sensitive analyzers based on phenylthiohydantoin (PTH) chemistry, it is now possible to obtain compositional data on samples separated by electrophoresis and blotted onto membranes (*UNIT 10.7*) such as are used for protein sequence analysis. This is especially helpful in cases where a protein sample fails to yield any sequence during gas-phase Edman degradation. Acid hydrolysis and analysis of sample blot pieces recovered from the sequencer sample compartment yield a rough estimate of the amount of protein present. If the amount is larger than that normally required for sequence analysis (10 pmol), it may then be concluded that the protein sample was N-terminally blocked.

CALCULATION OF THE AVERAGE COMPOSITION

Table 3.2.1 presents the results of analysis of a 24-kDa recombinant protein. The initial sample volume was 1.0 ml and was expected to contain ~4 mg total protein. A 30-µl aliquot was taken for analysis and was estimated to contain 120 µg, or ~5 nmol. The aliquot was transferred to a glass tube for hydrolysis and dried by application of a vacuum at room temperature. One milliliter of 6 M HCl was added, the tube was cooled in a dry ice bath, then a vacuum was applied and the tube sealed by heating in a gas flame. Hydrolysis was conducted for 24 hr at 124°C, followed by removal of the HCl under vacuum. The sample was taken up in 1.0 ml application buffer, 50 nmol norleucine was added as an internal standard, and a 50-µl aliquot was analyzed by ion-exchange chromatography on a Beckman 6300 system with ninhydrin detection.

1. The total number of nanomoles of amino acids detected is first calculated by summing the second column of Table 3.2.1, yielding 1446 nmol.

2. The total nanomoles is then divided by the total number of expected amino acids (based on protein sequence; in this case, 212) to obtain the average number of nanomoles per residue; 1446 nanomoles/212 residues = 6.82 nmol per residue. If the total number of amino acids is not known, it may be estimated by dividing the mol. wt. of the protein by 110.

3. The number of nanomoles of each amino acid measured is divided by the average number of nanomoles per residue (step 2) to give the ratio[1] observed for that amino acid. For Asp, 155 ÷ 6.82 = 22.7.

The resulting ratios are given in the fourth column and may be compared to the ratio expected listed in the third column. For most

ECU IAS/CERIAS - PURDUE UNIVERSITY - NOT FOR DISTRIBUTION

Table 3.2.1 Data Used to Calculate Amino Acid Composition for a 24-kDa Protein

Amino acid	Nanomoles measured	Ratio expected[a]	Ratio[1] observed	Ratio[2] observed
Asp	155	23	22.7	21.8
Thr	78	13	11.4	10.9
Ser	63	13	9.2	8.9
Glu	153	21	22.4	21.5
Pro	52	7	7.6	7.3
Gly	157	22	23.0	22.1
Ala	90	12	13.1	12.6
Val	159	23	23.3	22.3
Met	24	5	3.5	3.4
Ile	88	13	12.9	12.3
Leu	138	19	20.2	19.4
Tyr	43	6	6.3	6.1
Phe	56	8	8.2	7.9
His	16	3	2.3	2.3
Lys	121	17	17.7	17.0
Arg	53	7	7.8	7.4

[a]Ratio expected based on amino acid sequence of protein.

of the amino acids, the ratio[1] observed agrees with the ratio expected.

The value of average nanomoles per residue calculated in step 2 is based on the analysis of a 30-µl sample; the total sample of 1 ml contains ~5.46 mg of protein (average moles/residue × grams/mole protein × 1000/30). This analysis utilized only 50 µl of the 1.0 ml final solution, so a sample size smaller than 3% of the original sample could have been used. Using an excess, however, provides the flexibility to repeat the determination if the first analysis gives peaks that are either too large or too small for accurate integration.

The ratio[1] observed for Ser, Thr, His, and Met are in this sample are ~30%, 13%, 24%, and 30% below the expected values, respectively. These results are typical for these four amino acids, but for different reasons. The Ser and Thr results are low because these amino acids undergo partial degradation during acid hydrolysis. Ser and Thr losses of 10% to 40% are normal. The lower-than-expected values for His and Met are an anomaly that is most likely caused by the relatively low copy number of those two amino acids in the protein. Due to these known problems, most investigators choose to delete the values for sensitive residues (Ser, Thr) as well as the values for those amino acids appearing in low amounts (in this case, His, and Met) and calculate the average number of nanomoles per residue independently of those amino acids. This gives the corrected ratios shown in the last column of Table 3.2.1 (ratio[2] observed). In this example, the recalculated values are changed by as much as 0.9 due to the higher average number of nanomoles per residue (7.11).

The deviation of ratio[1] observed from the expected ratio ranges from 1% to 32% for the various amino acids, with an average error of 8.8%. When the high variation for Ser (32%), Met (32%), and His (24%) are removed from consideration, the average deviation of ratio[2] observed from expected ratio drops to 4.1%. By standard criteria, this sample would be judged to be pure; of course, this result needs to be combined with electrophoretic separation (Chapter 10), chromatographic purification (Chapter 8), and perhaps mass spectroscopic analyses before a final assessment of the purity of the sample can be made.

Key Reference

Ozols, J. 1990. Amino acid analysis. *Methods Enzymol.* 182:587-601.

Excellent discussion of sample preparation issues and hydrolysis methods.

Contributed by Ben Dunn
University of Florida
Gainesville, Florida

BUJARD-FETRAL UNIVERSITY LIBRARY

In Vitro Radiolabeling of Peptides and Proteins

Radiolabeling of peptides or proteins is often performed to enhance the sensitivity of detection, to quantitate the binding of peptides to other molecules, or for radioimmunoassays. The choice of radioisotope depends in part on the amino acid residue(s) to be modified (Table 3.3.1). The most common and straightforward labeling strategy is radioiodination with ^{125}I because of its relative ease of use, the high resulting specific radioactivities (curies per millimole), and the convenience of detecting γ emissions. Radioiodination may also be performed with ^{131}I, which has a shorter half-life (8 days versus 60 days for ^{125}I), giving higher specific radioactivities but limiting the useful lifetimes of the labeled products; in addition, ^{131}I has higher-energy γ emissions, necessitating more shielding to protect workers.

Iodination at tyrosine or histidine residues using Iodo-Beads (Basic Protocol 1) is the most convenient method, but chloramine T or Iodogen (Alternate Protocol 1) may give higher recoveries when working with small quantities of peptide, and lactoperoxidase (Alternate Protocol 2) results in less oxidative damage and lower specific radioactivities, and is commonly used when labeling whole cells. Modifications for achieving equimolar or higher (stoichiometric) rather than trace iodination are presented in Support Protocol 1, and the optional separation of unlabeled peptide from iodinated products by HPLC is described in Support Protocol 2. For peptides that do not contain tyrosine or histidine, or in which labeling of these amino acid residues interferes with biological activity, ^{125}I or ^{131}I may be added onto primary amino groups (lysine residues and the peptide N-terminus) using Bolton-Hunter reagent (Basic Protocol 2).

As an alternative to iodination, peptides or proteins can be labeled with ^{14}C or ^{3}H by acetylation of primary amino groups using anhydride (Basic Protocol 3), by reductive alkylation of primary amino groups using aldehyde and borohydride (Alternate Protocol 3), or by alkylation of cysteine sulfhydryl groups (Alternate Protocol 4). Figure 3.3.1 summarizes these radiolabeling strategies as well as the basic iodination reactions. Two methods are available to label a peptide without altering its structure: ^{3}H labeling by catalytic reduction of (nonradioactive) ^{127}I-labeled peptide (Basic Protocol 4) and de novo peptide synthesis using a radiolabeled form of any amino acid (Basic Protocol 5). However, these methods tend to be more expensive and laborious. Finally, selective

Table 3.3.1 Peptide Labeling Strategies

Amino acid residue	Label	Method	Reference
Tyrosine	^{125}I, ^{131}I	Chloramine T, Iodo-Beads	Greenwood et al. (1963)
Histidine, tyrosine	^{125}I, ^{131}I	Chloramine T, Iodo-Beads	Wolff and Covelli (1969)
Tyrosine	^{125}I, ^{131}I	Lactoperoxidase	Marchalonis (1969)
Lysine, N-terminus	^{125}I, ^{131}I	Bolton-Hunter	Bolton and Hunter (1973)
Lysine, N-terminus	^{14}C, ^{3}H	Anhydride	Fraenkel-Conrat and Colloms (1967)
Lysine, N-terminus	^{14}C, ^{3}H	Aldehyde/borohydride	Tack and Wilder (1981)
Cysteine	^{14}C, ^{3}H	Iodoacetic acid	Chersi et al. (1988)
Tyrosine	^{3}H	Reduction of ^{127}I	Amersham or Du Pont NEN literature
Any residue	^{14}C, ^{3}H	Peptide synthesis	Stewart and Young (1984)
Carbohydrate groups	^{3}H	Periodate/borohydride	Gahmberg and Andersson (1977)

Copyright © 1995 by John Wiley & Sons, Inc.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

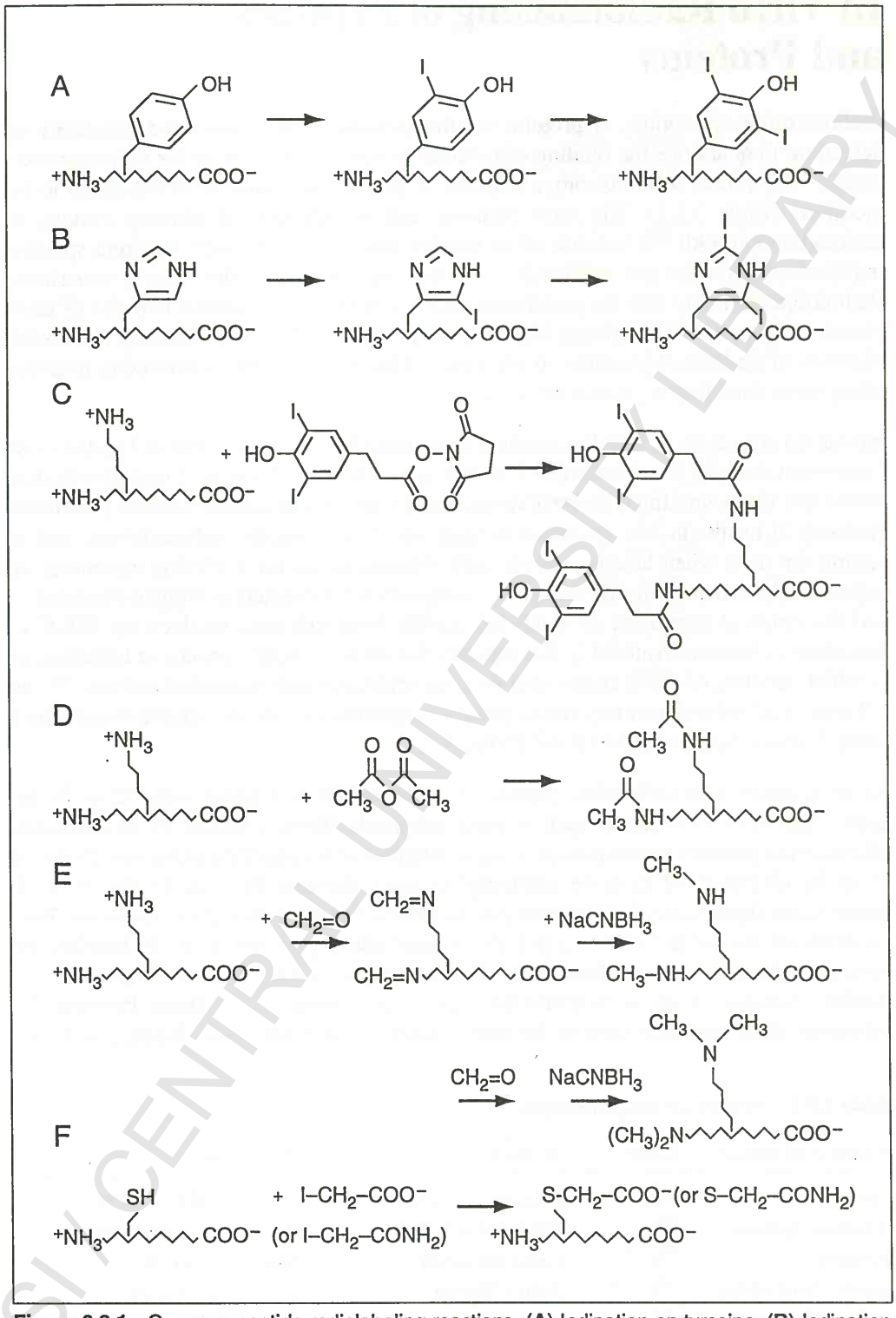


Figure 3.3.1 Common peptide radiolabeling reactions. **(A)** Iodination on tyrosine. **(B)** Iodination on histidine. **(C)** Bolton-Hunter iodination on lysine and N-terminus. **(D)** Acetylation on lysine and N-terminus. **(E)** Reductive alkylation on lysine and N-terminus. **(F)** Alkylation on cysteine.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

labeling at the N-terminus of a peptide can be performed during peptide synthesis (Alternate Protocol 5).

All peptide-labeling procedures described in this unit are also applicable to larger proteins; however, with larger proteins there is increased potential for idiosyncratic results because of tertiary structure and for undesired side reactions because of sensitive amino acid residues (e.g., tryptophan, methionine, and cysteine in the case of oxidative iodinations). To some extent these problems may be minimized through the use of milder labeling procedures (e.g., lactoperoxidase-catalyzed iodination).

NOTE: Various methods for separating radiolabeling agents (e.g., free iodide) from labeled peptide products are described in Basic Protocol 1 and in *UNITS 8.2 & 8.3*, and any of these methods may be used in conjunction with any of the radiolabeling protocols in this unit. The choice of separation method should depend primarily on the properties of the peptide or protein (see Critical Parameters, discussion of iodination work-up).

CAUTION: When working with ^{125}I, ^{131}I, ^{14}C, ^{3}H, or other radiochemicals, always observe appropriate safety precautions, including the use of gloves (consider use of a glove box with a charcoal filter, particularly for >5 mCi iodide), shielding, dosimetry monitoring, and proper disposal of radioactive waste. Consult a radiation safety specialist to ensure that all procedures are safe and in accordance with Nuclear Regulatory Commission guidelines. Oxidation converts iodide to a volatile and hazardous form; peptide-bound iodine is not volatile, but emits γ radiation and remains a potential hazard. Be sure to perform all steps that may give rise to volatile iodine in an externally vented hood. See *APPENDIX 2B* for additional guidelines.

IODINATION AT TYROSINE OR HISTIDINE RESIDUES USING IODO-BEADS

BASIC PROTOCOL 1

Iodo-Beads are nonporous polystyrene beads with an immobilized oxidizing agent (*N*-chlorobenzenesulfonamide) capable of converting $Na^{125}I$ (or $Na^{131}I$) to its reactive form (Markwell, 1982). Iodination occurs by electrophilic addition to tyrosine at the two positions *ortho* to the hydroxyl group, and on histidine at two imidazole carbon positions (Means and Feeney, 1971; Fig. 3.3.1). After reaction of a peptide with oxidized iodide, the mixture is transferred to a solid-phase extraction device (e.g., a Sep-Pak cartridge) to remove free iodide from the labeled peptide.

Materials

Iodo-Beads (Pierce)
0.1 M sodium phosphate, pH 6.0 *or* pH 8.5
Carrier-free sodium iodide-125 ($Na^{125}I$; Amersham, Du Pont NEN, or ICN Biomedicals) or sodium iodide-131 ($Na^{131}I$)
0.1 M sodium hydroxide
Peptide, ideally dissolved in ≤0.4 ml of water or 0.1 M sodium phosphate
Methanol
HPLC solvent A: 0.1% (v/v) trifluoroacetic acid (TFA) in water
HPLC solvent B: 0.1% (v/v) TFA in acetonitrile

Filter paper
100-µl syringe with beveled tip, or hot pipettor (i.e., dedicated to radioactive use)
Luer-tipped syringes (1-, 5-, and 20-ml) and one 23-G needle
Solid-phase extraction device (e.g., Sep-Pak or Sep-Pak Plus C18 cartridge, Waters)
Polypropylene or other plastic tubes (12 × 75 mm)

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

1. Wash Iodo-Beads with 1 ml of 0.1 M sodium phosphate, pH 6.0 (for labeling primarily on tyrosine) or pH 8.5 (for labeling on histidine and tyrosine). Dry on filter paper and add to 0.5 ml of same buffer in capped microcentrifuge tube or small reaction vial.

   For trace labeling use two Iodo-Beads; for stoichiometric labeling more may be needed (see Support Protocol 1).

   Iodo-Beads are compatible with most salts (including azide), chaotropic agents, and detergents, but not with reducing agents or organic solvents that dissolve polystyrene such as dimethyl sulfoxide (DMSO) and dimethylformamide (DMF). Iodo-Beads must be stored in the presence of desiccant, and have a manufacturer-suggested shelf life of 1 year.

2. Add $Na^{125}I$ (or $Na^{131}I$) to Iodo-Beads in the reaction tube using a 100-µl syringe with beveled tip or a hot pipettor. Cap tube and wait 5 min at room temperature, agitating the mixture occasionally. Rinse syringe with 0.1 M sodium hydroxide (^{125}I is less volatile at high pH) and then with water before storing.

   The amount of radioactive iodide depends largely on the desired specific radioactivity, subject to safety and cost limitations; 1 to 5 mCi is typical. Use a fresh batch to minimize loss of radioisotope as volatile molecular iodine. $Na^{125}I$ is supplied in various concentrations of NaOH; the more dilute NaOH solutions are preferable since they are less likely to alter the pH of the reaction.

3. Add peptide to the reaction. Cap tube and wait 30 min, agitating occasionally.

   The amount of peptide is inversely related to the final specific radioactivity, unless iodinated products are separated from unlabeled peptide (see Support Protocol 2); 0.05 to 2.0 mg is typical. Reaction time can be 45 min or longer if desired, or just 5 to 10 min if oxidation-sensitive residues in the peptide (tryptophan, methionine, cysteine) are of concern, and should be optimized empirically if needed.

4. During the iodination reaction (or before if more convenient), prepare a Sep-Pak C18 cartridge by attaching a syringe and passing 5 ml methanol through the cartridge, followed by 5 ml water and 10 to 20 ml HPLC solvent A. Maintain a flow rate of ~2 ml/min, do not reverse the orientation of the cartridge when changing solvents, and avoid introducing air bubbles or allowing the cartridge to dry out.

   Although this Sep-Pak method offers the most efficient separation of unbound radioactive iodide from the peptide, anion-exchange (UNIT 8.2) or gel-filtration (UNIT 8.3) chromatography may be used as alternatives, especially for larger peptides at risk of denaturation under the conditions used in steps 4 to 8 (see Alternate Protocol 1, steps 1 and 4). Dialysis is another alternative for large peptides, but takes much longer.

5. Remove the mixture from the reaction tube using a 1-ml syringe fitted with a 23-G needle. Carefully remove the needle and slowly load mixture onto the preconditioned Sep-Pak cartridge, collecting the eluate into a radioactive liquid waste container.

6. Wash Sep-Pak cartridge with 20 ml HPLC solvent A, collecting the eluate as radioactive waste. Neutralize waste with 5 ml of 0.1 M sodium hydroxide.

   Unbound radioactive iodide will pass through the Sep-Pak cartridge under these conditions, while peptide will remain in the cartridge.

7. Elute two fractions from the Sep-Pak cartridge into 12 × 75–mm plastic tubes, the first using 4 ml of a 1:1 mixture of HPLC solvent A and HPLC solvent B, and the second using 4 ml HPLC solvent B.

   Usually the first tube contains most of the labeled peptide, but a two-step elution improves overall recovery. Because unbound iodide is now removed, the tubes may be transferred out of the glove box or other iodination facility, as long as appropriate shielding is maintained.

PGU IASRI CENTRAL UNIVERSITY LIBRARY

8. Dry the tubes in a Speedvac evaporator. Redissolve in water or desired buffer and pool the two fractions.

9. To determine the specific radioactivity of the labeled peptide, estimate peptide recovery or analyze an aliquot of the sample for peptide concentration (e.g., by a colorimetric assay) and simultaneously determine cpm (in a γ-radiation counter, possibly after serial dilutions in the presence of a carrier protein to reach a level of radioactivity within the linear counting range of the γ counter).

   Specific radioactivity (in cpm/μg) = measured radioactivity (in cpm/μl) ÷ peptide concentration (in μg/μl). To convert to millicuries, use 1 mCi = 2.22 × 10^9 dpm and cpm = dpm × counting efficiency of γ counter (often 50% to 75%, depending on the instrument).

   Alternatively, if iodinated products are separated from unlabeled peptide by HPLC and the molar ratio of incorporated iodide to peptide is known (or estimated), the specific radioactivity of the peptide may be calculated from that of carrier-free ^{125}I (or ^{131}I; see Support Protocol 2).

IODINATION AT TYROSINE OR HISTIDINE RESIDUES USING CHLORAMINE T OR IODOGEN

ALTERNATE PROTOCOL 1

Instead of using Iodo-Beads, $Na^{125}I$ can be oxidized to its reactive form using soluble chloramine T (*N*-chloro-4-methylbenzenesulfonamide sodium salt; Greenwood et al., 1963), the solid-phase reagent Iodogen (1,3,4,5-tetrachloro-3α,6α-diphenylglycoluril; Fraker and Speck, 1978), or the enzyme lactoperoxidase (see Alternate Protocol 2). Although slightly less convenient than Iodo-Beads, these methods may give higher recoveries when working with small quantities of peptide (<0.05 mg), which can adsorb to Iodo-Beads. The reagents are inexpensive and stable, and with Iodogen or lactoperoxidase, expose the peptide to relatively mild reaction conditions.

***Additional Materials** (also see Basic Protocol 1)*

Anion-exchange resin, e.g., Dowex 1-X8, chloride form, 100 to 200 mesh (Bio-Rad)
BSA solution: 1.0 mg/ml BSA in PBS, with 0.02% sodium azide (NaN_3)
1.0 mg/ml chloramine T in 0.1 M sodium phosphate, prepared immediately before use
Iodogen (Pierce), dichloromethane, and glass tubes or vials (optional alternate to chloramine T)
Saturated solution of tyrosine in water (~0.4 mg/ml at 25°C)
2.5 mg/ml sodium metabisulfite ($Na_2S_2O_5$) in 0.1 M sodium phosphate, prepared immediately before use (optional alternate to tyrosine solution)
Pasteur pipet or 1-ml syringe with glass wool plug inserted at bottom
Gel-filtration column, e.g., Sephadex G-10 or G-25 (Pharmacia Biotech), or Excellulose Desalting Column (Pierce; optional alternate to anion-exchange column)

1. Suspend Dowex resin in water and pour into plugged Pasteur pipet. Wash with ≥20 ml BSA solution. Alternatively, equilibrate a gel-filtration column with BSA solution. Avoid introducing air bubbles into the column and do not allow column to dry out.

   Because short peptides (<20 residues) are difficult to separate completely from unbound radioactive iodide by gel-filtration columns, an anion-exchange column may be preferred. Highly negatively charged (e.g., phosphorylated) peptides and very short peptides (<10 residues) may be retained on an anion-exchange column and should therefore be desalted by solid-phase extraction (see Basic Protocol 1). The desalting method should be established in advance.

BCU LASS CENTRAL UNIVERSITY LIBRARY

2a. *Chloramine T procedure:* Place peptide in a small tube and buffer with 0.1 M sodium phosphate, pH 6.0 or 8.5 (see Basic Protocol 1, step 1); bring the total volume to 100 to 300 µl. Add $Na^{125}I$ (or $Na^{131}I$), followed by 10 µl chloramine T (freshly prepared), and mix. Wait 1 min for reaction to occur.

2b. *Iodogen procedure:* Add sodium iodide to the peptide (as described in step 2a) in a glass tube previously coated with the solid-phase reagent Iodogen by dissolving the solid in dichloromethane and blowing off the solvent under a stream of dry nitrogen. Wait 10 min at room temperature.

Iodogen-coated tubes may be stored for many months at room temperature in a desiccator, protected from light. Because the oxiding agent is water insoluble, peptide damage is kept to a minimum.

3. Add 50 µl saturated tyrosine solution to quench the reaction.

Alternatively, the reaction may be terminated by adding 10 µl sodium metabisulfite, although this reducing agent is potentially more damaging to peptides.

4. Load mixture onto a Dowex or gel-filtration column and collect 0.5-ml fractions in 12 × 75–mm plastic tubes, using BSA solution for elution. Monitor fractions with a hand-held γ counter, or count 1-µl aliquots of each fraction in a γ counter to locate the peptide-containing fractions.

Radiolabeled peptide should pass through the column; unreacted iodide is retained longer and need not be eluted. Dispose of the entire column as radioactive sharps waste.

ALTERNATE PROTOCOL 2

IODINATION AT TYROSINE OR HISTIDINE RESIDUES USING LACTOPEROXIDASE

Although the degree of oxidative damage to peptides undergoing iodination is difficult to predict, the lactoperoxidase method results in less oxidative damage than chemical oxidative procedures (Marchalonis, 1969). Iodination is confined to tyrosine. Specific radioactivities are likely to be lower by this method. Because lactoperoxidase (mass 77,500 Da) will be iodinated during the reaction, it should be used only if the labeled peptide or protein of interest can be separated from labeled lactoperoxidase by gel filtration (*UNIT 8.3*), ion-exchange chromatography (*UNIT 8.2*), or some other means. Alternatively, it may be possible to use lactoperoxidase coupled to beads, or biotinylated lactoperoxidase and avidin-coated beads to remove the labeled enzyme (Sigma). Because the enzyme does not gain access to the cytoplasm, lactoperoxidase-catalyzed iodination is useful for selectively radiolabeling membrane proteins on whole cells, provided iodide concentrations are kept low and the proportion of dead cells is very small (Marchalonis et al., 1971).

Additional Materials (also see Basic Protocol 1)

0.2 mg/ml lactoperoxidase in PBS
0.01% hydrogen peroxide (H_2O_2) in PBS (diluted from 30% stock immediately before use)
2 U/ml glucose oxidase in PBS and 0.1 M D-glucose in PBS (optional alternate to H_2O_2)

1. Place peptide in a tube with PBS, ideally at ≥1.0 mg/ml in a small volume.
2. Add 10 µl of 0.2 mg/ml lactoperoxidase.

Azide should not be present in the reaction as it inhibits lactoperoxidase.

BCU IASL / CENTRAL UNIVERSITY LIBRARY

3. Add $Na^{125}I$ (or $Na^{131}I$), cap tube, and mix.

4. Add 10 µl of H_2O_2. Wait 10 min at room temperature.

   The H_2O_2 concentration is critical and may need fine-tuning for best results. Alternatively, use a coupled glucose/glucose oxidase system to generate H_2O_2 in situ, i.e., treat with 10 µl of glucose oxidase and 10 µl of glucose for 10 min at 4°C (Hubbard and Cohn, 1975).

5. Terminate the reaction and remove unbound radioactive iodide (see Alternate Protocol 1, steps 1, 3, and 4).

EQUIMOLAR IODINATION TO GIVE HIGH YIELDS OF PRODUCT

SUPPORT PROTOCOL 1

Iodination of microgram to milligram amounts of a peptide using low millicurie amounts of carrier-free $Na^{125}I$ (or $Na^{131}I$), such as those suggested in Basic Protocol 1 and Alternate Protocols 1 and 2, results in trace labeling (iodination of only a small fraction of the peptide molecules). For example, in a reaction between 1 mCi of ^{125}I (~600 pmol) and 500 µg of a peptide of mass 1000 Da (500 nmol) or a protein of mass 100,000 Da (5 nmol, or ~200 nmol of tyrosine residues on average), there is at most only enough ^{125}I to label well under 1% of the tyrosines. For proteins, trace labeling is often advantageous as it minimizes the chance that the protein's activity will be lost (Eisen and Keston, 1949; Hunter and Greenwood, 1962); moreover, a sizable fraction of protein molecules will have one iodine atom on at least one of its tyrosines. However, for short peptides over 99% of the molecules remain unmodified, greatly diluting the specific radioactivities that are possible and complicating the interpretation of experiments in which labeled and unlabeled peptides differ significantly in their activities (see Support Protocol 2).

To label a higher proportion of a peptide, nonradioactive $Na^{127}I$ ("carrier") can be added to ^{125}I, giving a higher yield of iodinated product and resulting in a mixture of chemically equivalent ^{125}I- and ^{127}I-labeled peptides. Perform all steps for iodination as described using Iodo-Beads (see Basic Protocol 1) or chloramine T or Iodogen (see Alternate Protocol 1), with two modifications. First, add 0.1 M $Na^{127}I$ in water to the $Na^{125}I$ prior to using $Na^{125}I$ in the indicated step. The amount of ^{127}I should be in moderate molar excess over the peptide (2- to 5-fold) to achieve high yields of iodinated products (70% to 95%). Further increasing the amount of ^{127}I unnecessarily diminishes the products' specific radioactivities: e.g., at a 4-fold molar excess of ^{127}I, only up to 25% of the ^{125}I added can be incorporated, at a 10-fold molar excess only 10% can be incorporated, etc. Second, a sufficient number of Iodo-Beads (or other oxidizing reagent) must be used to oxidize the total iodide (essentially ^{127}I). Use the manufacturer-determined oxidizing capacity of 0.55 ± 0.05 µmol per Iodo-Bead. Note that Iodo-Beads rapidly develop a yellow-brown color when stoichiometric amounts of iodide are added; failure to observe this color indicates deterioration of the beads.

For maximum specific radioactivities, label a small amount of peptide (e.g., 30 µg) with up to an equimolar amount of ^{125}I (e.g., 40 mCi) and separate the unlabeled peptide from iodinated products by HPLC (see Support Protocol 2).

It is also possible to iodinate with ^{127}I only, in which case the products will be nonradioactive but chemically equivalent to the corresponding radiolabeled products. This approach is useful for practicing the iodination procedure and for assessing the effects of iodination on a peptide's activity (Tsomides and Eisen, 1993), as well as for preparing an ^{127}I-labeled peptide to be reductively tritiated (see Basic Protocol 4).

BOSTON UNIVERSITY LIBRARY

SUPPORT PROTOCOL 2

SEPARATION OF UNLABELED PEPTIDE FROM IODINATED PRODUCTS BY HPLC

Whether trace iodinated with ^{125}I only or stoichiometrically iodinated with a mixture of ^{125}I and ^{127}I, some peptide will remain unlabeled, and it may be desirable to separate this unlabeled component from the iodinated products. To obtain chemically homogeneous radiolabeled peptides, iodinated products can be resolved from each other and from unlabeled peptide by reversed-phase HPLC. (This optional procedure is distinct from the obligatory separation of peptide from unbound radioactive iodide by solid-phase extraction, anion-exchange chromatography, gel filtration, exhaustive dialysis, or other means as detailed in steps 4 to 8 of Basic Protocol 1.) Because iodine is a bulky, hydrophobic atom relative to hydrogen, the retention times of iodinated peptides will be greater than those of their unlabeled counterparts on reversed-phase HPLC columns (e.g., C18 for short peptides, C4 for longer or more hydrophobic peptides).

In most iodinations, more than one product will be generated, even if the peptide has only a single tyrosine (see Fig. 3.3.1). Because the pK_a of monoiodotyrosine is lower than that of tyrosine (Edelhoch, 1962), electrophilic addition of the second iodine atom may be accelerated relative to the first; however, this reaction will not go to completion even when iodide is present in molar excess over the peptide. When more than two or three iodinatable residues (tyrosine and histidine) are present (as in a protein), the labeled products may be too complex to allow for individual resolution by HPLC, but unlabeled peptide can still be removed from the radiolabeled mixture. Even peptides as long as 50 to 100 residues can be separated from their iodinated derivatives by reversed-phase HPLC.

For peptides with a single iodinatable residue, the two products are usually easily separated by HPLC and can readily be identified by their specific radioactivities, which are in the ratio 1:2 for monoiodinated and diiodinated peptide, respectively. Figure 3.3.2 shows the result of stoichiometrically labeling 2.0 mg of an eight-residue peptide with 20

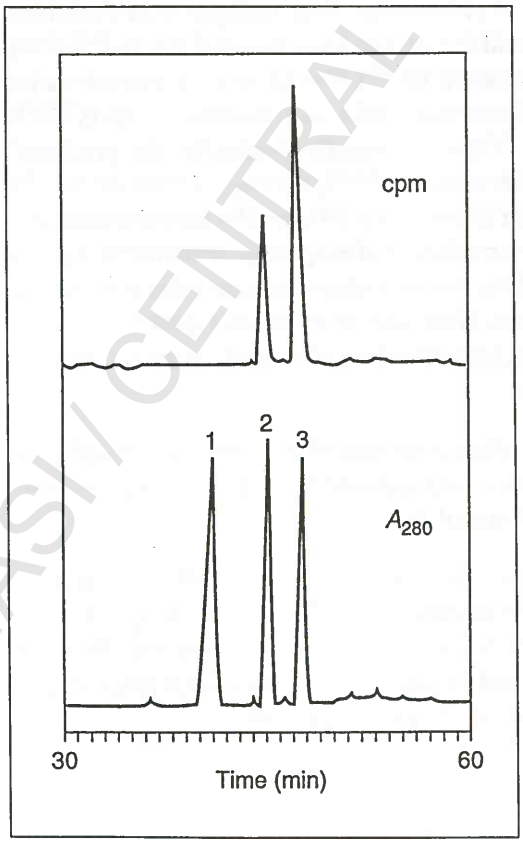


Figure 3.3.2 Stoichiometric iodination of the eight-residue peptide LSPFPYDL and HPLC separation of the unlabeled starting material (peak 1) from monoiodinated (peak 2) and diiodinated (peak 3) products. Two milligrams of peptide were iodinated with 20 mCi $Na^{125}I$ and 5 µmol $Na^{127}I$ at pH 7.0 using six Iodo-Beads, injected on a Vydac C4 semipreparative column, and eluted with an acetonitrile gradient of 1%/min at a flow rate of 3 ml/min. The specific radioactivities of the three peaks were $<10^5$ cpm/µg, 2.1×10^7 cpm/µg, and 4.6×10^7 cpm/µg, respectively.

PGU IASI / CENTRAL UNIVERSITY LIBRARY

mCi $Na^{125}I$ mixed with 5 μmol $Na^{127}I$ and separating the unreacted, monoiodinated, and diiodinated products by HPLC. If Edman degradation is available, labeled products may be characterized directly by detecting the monoiodinated and diiodinated derivatives of tyrosine and/or histidine (Tsomides and Eisen, 1993). Note that peptides containing tryptophan, methionine, or cysteine residues may undergo oxidation during iodination. The resulting oxidized products (both radiolabeled and unlabeled) may appear as separate peaks on HPLC, complicating the identification of products.

B.C.U. „M. Eminescu" IAŞI
CHIMIE

When carrying out stoichiometric iodination with a mixture of ^{125}I and ^{127}I, specific radioactivities of the product mixture or of individual HPLC-purified products can be measured as described for the Iodo-Beads procedure (see Basic Protocol 1, step 9). However, when trace labeling with ^{125}I only and removing unlabeled peptide by HPLC, specific radioactivities may simply be calculated as the specific radioactivity of carrier-free ^{125}I (~2000 Ci/mmol) multiplied by the molar ratio of ^{125}I to peptide (e.g., 1 for a monoiodinated derivative, 2 for a diiodinated derivative, etc.). When a mixture of ^{125}I-labeled products is separated from unlabeled peptide by HPLC but not resolved into its individual components, the same calculation may be applied by taking the average molar ratio of ^{125}I to peptide to be the number of iodinatable residues. For example, the specific radioactivity of a peptide of mass 1000 Da with two tyrosine residues that is labeled with ^{125}I and separated from unlabeled peptide by HPLC is $\sim 5 \times 10^9$ cpm/μg, assuming an average of two ^{125}I per labeled peptide.

NOTE: Although HPLC fractions can be analyzed for radioactivity by transferring them to a γ counter, it is preferable to use an on-line radioisotope detector connected in series with an ultraviolet absorbance detector (see Fig. 3.3.2).

IODINATION AT LYSINE RESIDUES OR N-TERMINUS USING BOLTON-HUNTER REAGENT

BASIC PROTOCOL 2

Many peptides do not contain tyrosine or histidine. Other peptides' function may be altered if crucial tyrosine or histidine residues are iodinated or if damage to oxidative-sensitive residues occurs during iodination. An alternate strategy is to label primary amino groups on lysine residues and/or at the N-terminus of the peptide using Bolton-Hunter reagent 3-(4-hydroxyphenyl)propionic acid *N*-hydroxysuccinimide ester (Bolton and Hunter, 1973). This reagent introduces a side chain similar to that of tyrosine (see Fig. 3.3.1). Of course, this method carries the potential drawback of altering peptide function if lysine residues (or the N-terminus) are critical for activity; unlike the methods above, peptide charge will be decreased by one unit for each site modified (at pH values below the amino group's pK_a value).

Because lysine's ε-amino group and the N-terminal α-amino group have different pK_a values (~10 vs. 8 to 9), it is possible, in principle, to achieve selective labeling by adjusting the pH (lower pH should favor the N-terminus). However, this is difficult in practice, and a mixture of products is to be expected from lysine-containing peptides unless the N-terminal residue is proline (secondary amine) or a protecting group is present on either lysine or the N-terminus.

^{125}I-labeled Bolton-Hunter reagent is commercially available in monoiodinated (~2000 Ci/mmol) or diiodinated (~4000 Ci/mmol) form. A cost-saving modification is to prepare ^{125}I-labeled Bolton-Hunter reagent from the less expensive unlabeled reagent by iodination using Iodo-Beads (see Basic Protocol 1) or chloramine T (see Alternate Protocol 1) followed by extraction of the product into anhydrous benzene. (In this case it is crucial to work quickly to minimize hydrolysis, and specific radioactivities will be lower unless unreacted Bolton-Hunter reagent is removed from the labeled product.) Alternatively, the

BCU IASI CENTRAL UNIVERSITY LIBRARY

peptide can first be modified with unlabeled Bolton-Hunter reagent and then radioiodinated using any of the methods previously described. Because ^{125}I-labeled Bolton-Hunter reagent suffers some degradation by competing hydrolysis during its reaction with peptide, this alternate approach gives a higher incorporation of radioactivity; however, it also results in the radiolabeling of any tyrosine or histidine residues present in the peptide, as well as exposure of the peptide to oxidative conditions.

Materials

0.1 M sodium borate, pH 8.5
0.1% (w/v) gelatin in 0.1 M sodium borate, pH 8.5
^{125}I-labeled Bolton-Hunter reagent (Amersham, Du Pont NEN, or ICN Biomedicals); or substitute unlabeled Bolton-Hunter reagent and iodinate later
Peptide, ideally ≥1 mg/ml in ≤100 µl of 0.1 M sodium borate, pH 8.5
1 M glycine in 0.1 M sodium borate, pH 8.5

Gel-filtration column, e.g., Sephadex G-10
Dry nitrogen tank and outlet tubing fitted with needle
Polypropylene or other plastic tubes (12 × 75 mm)

1. Equilibrate a gel filtration column with 0.1% gelatin/0.1 M sodium borate, pH 8.5, to prevent loss of radiolabeled peptide by nonspecific adsorption.

   BSA is not recommended as a carrier protein in place of gelatin because albumin binds to Bolton-Hunter reagent. If gelatin is undesired in the final preparation (e.g., iodination is to be performed later), condition the column and then reequilibrate in buffer of choice. Alternatively, use a Sep-Pak cartridge instead of gel filtration to separate by-products and salts from the labeled peptide (see Basic Protocol 1, steps 4 to 8).

2. Immediately prior to use, evaporate ^{125}I-labeled Bolton-Hunter reagent to dryness by inserting a needle through the rubber septum and admitting a gentle stream of dry nitrogen, using the manufacturer-supplied charcoal trap as a vent to contain any released volatile radioiodine. Do not splash solution onto upper parts of the vial.

   If using unlabeled Bolton-Hunter reagent, prepare a stock solution in anhydrous organic solvent and dilute into borate buffer at the time of reaction to achieve a 3- to 5-fold molar excess over peptide. Unlabeled Bolton-Hunter reagent is also available as a water-soluble (sulfo) derivative (at higher cost), obviating the need for organic solvent and offering an advantage for labeling cell surface proteins because the sulfo derivative is membrane impermeable.

3. Chill peptide in borate buffer to 0°C on ice. Add peptide to vial containing dried ^{125}I-labeled Bolton-Hunter reagent (or freshly dissolved unlabeled Bolton-Hunter reagent) and wait 15 to 30 min at 0°C, agitating the reaction occasionally.

   Incorporation is greatest when using high peptide concentrations. Phosphate or bicarbonate buffers, pH 8.5, may be used instead of borate, but avoid amine-containing buffers (ammonium, glycine, Tris) as well as azides and thiols. Being an N-hydroxysuccinimide ester, Bolton-Hunter reagent is sensitive to hydrolysis, with a half-life of a few hours at pH 7.0 and only ~10 min at pH 8.5; however, its reaction with amino groups is greatly accelerated at the higher pH since more amine is in the deprotonated (active) form.

4. Add 100 µl of 1 M glycine in 0.1 M sodium borate, pH 8.5, to quench the reaction and wait 5 min, or use a corresponding amount of Tris or any other amine for quenching.

   Longer incubations at 0°C (>4 hr) may improve yields and obviate the need for quenching by consuming all of the reagent.

PGU IASI LIBRARY

5. Load mixture onto gel filtration column and collect 0.5-ml fractions in 12 × 75–mm plastic tubes, using borate (or other) buffer for elution. Monitor fractions with a hand-held γ counter, or count aliquots of each fraction in a γ counter to locate the peptide-containing fractions.
6. Resolve products by HPLC if desired and determine specific radioactivities (see Support Protocol 2).

^{14}C OR ^{3}H LABELING AT LYSINE RESIDUES OR N-TERMINUS BY ACETYLATION USING ANHYDRIDE

BASIC PROTOCOL 3

There are two main advantages of labeling with ^{14}C or ^{3}H instead of ^{125}I. First, for peptides that are unacceptably altered as a result of iodination, different sites may be targeted for modification (i.e., lysine or cysteine), or the peptide may even be labeled without changing its structure by replacing carbon or hydrogen atoms with the corresponding radioisotope (see Basic Protocol 4 and Basic Protocol 5). Second, the half-lives of ^{14}C and ^{3}H are much longer than that of ^{125}I (5760 and 12.26 years, respectively, vs. 60 days), so labeling need not be done often. However, the specific radioactivities of peptides labeled with ^{14}C or ^{3}H are generally correspondingly lower, and the labeling techniques are not as widely applied as radioiodination. Also, counting β emissions involves the use of a liquid scintillation cocktail and is therefore slightly more cumbersome than counting γ emissions.

Acetylation with [^{14}C]- or [^{3}H]anhydride is a simple reaction that will modify amino groups (lysine and/or the N-terminus of a peptide; Fig. 3.3.1; Fraenkel-Conrat and Colloms, 1967). The most common side reaction is O-acetylation of tyrosine, which is reversible under mild alkali conditions (Means and Feeney, 1971).

CAUTION: All procedures should be carried out in an externally vented hood.

Materials

Saturated sodium acetate solution
Peptide, ideally ≥1 mg/ml in ≤100 µl H_2O or sodium acetate
[^{14}C]- or [^{3}H]acetic anhydride (Amersham, Du Pont NEN, ICN Biomedicals, or Sigma)
Acetic anhydride, unlabeled (optional)

1. Add 1 vol of saturated sodium acetate solution to the peptide (unless already in this buffer) and cool to 0°C on ice.
2. Add [^{14}C]- or [^{3}H]acetic anhydride, with or without 2 molar equivalents of additional unlabeled acetic anhydride, in four equal installments at 10- to 15-min intervals.

   Because the anhydride is rapidly hydrolyzed, addition over several installments is the preferred method. The use of excess unlabeled anhydride increases the extent of reaction but dilutes the specific radioactivity of the products (analogous to stoichiometric vs. trace iodination; see Support Protocol 1).

3. Remove by-products by gel filtration (*UNIT 8.3*), dialysis (*APPENDIX 3B*), or solid-phase extraction (see Basic Protocol 1, steps 4 to 8), and count fractions in an appropriate liquid scintillation cocktail.

   Resolution of individual products may be possible by HPLC (see Support Protocol 2).

BCU ASTON CENTRAL UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 3

^{14}C OR ^{3}H LABELING AT LYSINE RESIDUES OR N-TERMINUS BY REDUCTIVE ALKYLATION

An alternative two-step procedure used to label peptides is treatment with an aldehyde followed by reduction of the resulting Schiff base with sodium cyanoborohydride (Fig. 3.3.1; Tack and Wilder, 1981; Jentoft and Dearborn, 1983). Most aldehydes give the monosubstituted alkylamine product, but formaldehyde goes on to give the dimethyl product. ^{14}C can be incorporated via the aldehyde, or ^{3}H can be added via labeled sodium cyanoborohydride. The latter gives higher specific radioactivities, reportedly ≥50 Ci/mmol. One advantage of this approach over acetylation is that the net charge on amino groups is not substantially altered (there is only a small change in pK_a value).

CAUTION: All procedures should be carried out in an externally vented hood.

Materials

0.2 M sodium borate, pH 9.0
Peptide, ideally ≥1 mg/ml in 0.1 to 1.0 ml of 0.2 M sodium borate, pH 9.0
[^{3}H]sodium borohydride (Amersham, Du Pont NEN, or ICN Biomedicals) or [^{3}H]sodium cyanoborohydride (Amersham)
0.01 M NaOH
37% (12.4 M) formaldehyde stock
[^{14}C]formaldehyde and unlabeled sodium cyanoborohydride (optional alternative for labeling with ^{14}C instead of ^{3}H)

1. Add an equal volume of 0.2 M sodium borate, pH 9.0, to the peptide (unless already in this buffer) and cool to 0°C on ice.

2. Dilute formaldehyde stock (12.4 M) to 0.1 to 0.2 M and add to peptide on ice to give a 2- to 6-fold molar excess over amino groups (e.g., 10 to 20 µl).

   More concentrated formaldehyde in this step may lead to peptide cross-linking.

   For ^{14}C labeling, use [^{14}C]formaldehyde in this step.

3. Immediately dissolve [^{3}H]sodium borohydride (or [^{3}H]sodium cyanoborohydride) in 0.01 M NaOH to a final concentration of 1.0 Ci/ml and add to peptide on ice for 10 min. Use a molar ratio of sodium borohydride to formaldehyde of between 0.25 and 0.40 (e.g., 30 to 50 µl).

   Sodium cyanoborohydride is a superior reagent for this purpose because it reduces the Schiff base formed initially between the peptide and formaldehyde without reducing disulfides or aldehydes. [^{3}H]sodium borohydride or [^{3}H]sodium cyanoborohydride should be stored frozen in small aliquots rather than frozen and thawed repeatedly.

   For ^{14}C labeling, use unlabeled sodium cyanoborohydride in this step.

4. Remove by-products by gel filtration (*UNIT 8.3*), dialysis (*APPENDIX 3B*), or solid-phase extraction (see Basic Protocol 1, steps 4 to 8), and count fractions in an appropriate liquid scintillation cocktail.

   Resolution of individual products may be possible by HPLC (see Support Protocol 2).

UCLA IASL - CHARLES E. YOUNG RESEARCH LIBRARY

^{14}C OR ^{3}H LABELING AT CYSTEINE RESIDUES USING IODOACETIC ACID OR IODOACETAMIDE

ALTERNATE PROTOCOL 4

To label cysteine residues, a peptide is reduced and allowed to react with either iodoacetic acid (to give the carboxymethyl derivative) or iodoacetamide (to give the carboxamidomethyl derivative, which is uncharged; see Fig. 3.3.1). For large peptides containing disulfide bonds, this procedure is often used to denature the peptide irreversibly. Carboxymethylcysteine content is readily determined by amino acid analysis following acid hydrolysis. [^{14}C]- or [^{3}H]iodoacetic acid and [^{14}C]iodoacetamide are commercially available for radiolabeling. Possible side reactions are alkylation of histidine, methionine, or lysine, each of which can usually be controlled by pH and other reaction conditions (Crestfield et al., 1963; Means and Feeney, 1971).

CAUTION: All procedures should be carried out in an externally vented hood.

Materials

0.5 M Tris·Cl/6 M guanidine·HCl/0.002 M disodium EDTA, pH 8.6
Peptide, ideally ≥1 mg/ml in 0.1 to 1.0 ml 0.5 M Tris·Cl/6 M guanidine·HCl/0.002 M disodium EDTA, pH 8.6
Dithiothreitol
1.0 M NaOH
[^{14}C]- or [^{3}H]iodoacetic acid or [^{14}C]iodoacetamide (Amersham, Du Pont NEN, or ICN Biomedicals)
Iodoacetic acid or iodoacetamide, unlabeled (optional)

1. Peptides containing disulfide bonds or significant tertiary structure require reduction and/or denaturation prior to alkylation of cysteine residues. For this, prepare a solution of peptide in 0.5 M Tris·Cl/6 M guanidine·HCl/0.002 M disodium EDTA (pH 8.6), incubate for 1 hr at 37°C, add dithiothreitol (50-fold molar excess over disulfide), and incubate overnight at 37°C.

2. Cool to room temperature and add desired amount of radiolabeled iodoacetic acid (or iodoacetamide) dissolved in 1/10 vol of 1.0 M NaOH for 30 min in the dark.

   Optionally, include additional unlabeled iodoacetic acid (or iodoacetamide) at a 2-fold molar excess over the dithiothreitol added to increase the extent of the reaction (analogous to stoichiometric vs. trace iodination; see Support Protocol 1).

3. Remove by-products by gel filtration (*UNIT 8.3*), dialysis (*APPENDIX 3B*), or solid-phase extraction (see Basic Protocol 1, steps 4 to 8), and count fractions in an appropriate liquid scintillation cocktail.

   Resolution of individual products may be possible by HPLC (see Support Protocol 2).

^{3}H LABELING AT TYROSINE RESIDUES BY REDUCTION OF ^{127}I-LABELED PEPTIDE

BASIC PROTOCOL 4

It is possible to label one or more tyrosine residues of a peptide stoichiometrically with ^{3}H to a relatively high specific radioactivity (e.g., 50 Ci/mmol). Since ^{3}H substitutes for hydrogen, the peptide's structure remains unaltered. However, this approach is more laborious and expensive than the previous methods, requiring the preparation of a ^{127}I-labeled peptide that is then catalytically reduced in the presence of tritium gas, leading to replacement of each ^{127}I (or other halogen atom) with ^{3}H.

The ^{127}I-labeled peptide precursor can be produced in two ways: by stoichiometric iodination with ^{127}I (see Basic Protocol 1, Support Protocol 1, and Support Protocol 2) or by incorporating a suitably protected ^{127}I-labeled tyrosine derivative (either mono-

BOULDER CENTER UNIVERSITY LIBRARY

iodinated or diiodinated) during de novo peptide synthesis (e.g., *N*-Boc-O-3-bromobenzyl-3,5-diiodo-L-tyrosine, available from Peninsula). The ^{127}I-labeled peptide is then submitted to a specialized facility for reductive tritiation (Amersham and Du Pont NEN offer this service).

BASIC PROTOCOL 5

INTRODUCTION OF LABELED AMINO ACID RESIDUES DURING PEPTIDE SYNTHESIS

The most specific and least disruptive labeling procedure is incorporation of a suitably protected radioactive amino acid into any desired position of a peptide during its solid-phase synthesis. However, because this requires an entire synthesis rather than use of just an aliquot of peptide, and because the protected derivatives of radioactive amino acids are not generally available and therefore must be prepared, it is not commonly done. Furthermore, introduction of a radioactive amino acid into a peptide synthesizer will be possible only if the operator of the instrument is prepared to work with radioactive samples. However, amino acids containing ^{14}C, ^{3}H, or ^{35}S (for methionine) are available, procedures for attaching various protecting groups to these amino acids are quite well established, and the result may be an abundant supply of a radiolabeled peptide unperturbed in structure and not obtainable by any other means. Specific radioactivities are typically two to three orders of magnitude lower than for radioiodinated peptides.

The simplest approach is to select a radioactive amino acid that requires protection during solid-phase peptide synthesis only on its α-amino group and not on its side chain (e.g., ^{14}C- or ^{3}H-labeled alanine, isoleucine, leucine, valine, or phenylalanine, all available from Amersham, Du Pont NEN, ICN Biomedicals, or Sigma). If classical *t*-butyloxycarbonyl (*t*-BOC) chemistry is to be used during peptide synthesis, a variety of procedures are available for attaching *t*-BOC to α-amino groups (Stewart and Young, 1984); one such procedure is outlined below for leucine.

Materials

L-Leucine, unlabeled
L-[^{14}C]- or L-[^{3}H]leucine, 1 to 5 mCi
Triethylamine
2-(*t*-butoxycarbonyloxyimino)-2-phenylacetonitrile (BOC-ON, Aldrich)
Dioxane
Diethyl ether, anhydrous
1.0 M HCl, 4°C
Separatory funnel

1. Add 131 mg L-leucine (1 mmol) and the desired number of millicuries L-[^{14}C]- or L-[^{3}H]leucine (supplied in water) to a 25-ml round-bottom flask.

   Other amino acids or their side chain protected derivatives can be substituted for leucine. Unlabeled amino acid is used because the amount of ^{14}C- or ^{3}H-labeled material is very small in molar terms (e.g., 2 mCi [^{3}H]leucine is <1 μmol); safety and cost considerations generally preclude the use of stoichiometric amounts of radiolabeled amino acid.

2. Add 210 μl triethylamine (1.5 mmol).

3. Add 271 mg BOC-ON (1.1 mmol).

4. Add a mixture of 600 μl dioxane and enough water to bring the total volume of water to 600 μl; e.g., if the radiolabeled amino acid is supplied in 400 μl water, add 600 μl dioxane and 200 μl water.

FGCU IASL / CENTRAL UNIVERSITY LIBRARY

5. Stir with a magnetic stir bar 3 hr at room temperature. The yellow, heterogeneous reaction will become homogeneous.

6. Add 2 ml water and 2 ml diethyl ether and transfer the reaction mixture to a separatory funnel.

   The aqueous (bottom) layer contains the triethylammonium salt of the t-BOC-amino acid product; the ether (upper) layer contains an oxime by-product from the reaction (bright yellow).

7. Drain the aqueous layer and wash it three times with an equal volume of diethyl ether in the separatory funnel to remove unwanted oxime.

8. Add a small volume of 4°C 1.0 M HCl to bring the aqueous layer to pH 2 and precipitate the *t*-BOC-amino acid product.

9. Add an equal volume of diethyl ether to the aqueous layer in the separatory funnel, shake, and drain off the aqueous layer. Save the ether layer and repeat ether extraction of the aqueous layer two more times. Pool the combined ether layers into an Erlenmeyer flask.

   The t-BOC-amino acid product now partitions into the ether layer after being converted by HCl from a salt to the free acid. Some t-BOC-amino acids (serine, threonine) remain soluble in water and are best extracted if the aqueous layer is first saturated with NaCl before adding ether.

10. Allow the product to crystallize for several hours on ice or overnight at 4°C. First blowing off some ether with a gentle stream of dry nitrogen gas may be helpful. Wash the white, crystalline product carefully with cold diethyl ether.

    The t-BOC-amino acid should be checked for nonreactivity of its amino group by a ninhydrin (Sarin et al., 1981) or other analytical test prior to using it in peptide synthesis.

SELECTIVE LABELING ON N-TERMINUS DURING PEPTIDE SYNTHESIS

ALTERNATE PROTOCOL 5

The N-terminus of a peptide that contains one or more lysine amino groups can be selectively labeled after the peptide has been synthesized by solid-phase methods but before amino acid side chains (including that of lysine) have been deprotected. For example, N-terminal acetylation of a newly synthesized protected peptide can be used to label the peptide with ^{14}C or ^{3}H without affecting lysine residues that might be important for activity (see Basic Protocol 3). Alternatively, addition of a Bolton-Hunter group allows subsequent iodination at this position (see Basic Protocol 2); tyrosine, if present, will also be iodinated unless radiolabeled Bolton-Hunter reagent is used.

Various other reagents can be used to modify a peptide at its N-terminus. For example, a dinitrophenyl (DNP) group can be attached specifically to the N-terminus of a protected resin-bound peptide using dinitrofluorobenzene (100 mol equivalents), *N,N*-diisopropylethylamine (20 mol equivalents), and dichloromethane as the solvent for the coupling reaction (2 hr at room temperature with shaking in the dark; T. Tsomides, unpub. observ.).

Specific details of the protocols to be used for N-terminal labeling reactions depend on the particular chemistries involved in the peptide's synthesis and cleavage. In particular, any label attached selectively to the N-terminus must be stable during the peptide side chain deprotection step (e.g., hydrogen fluoride or trifluoroacetic acid). In the example above, DNP is stable during deprotection with trifluoromethane sulfonic acid.

BRUNEL CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

Detection of minute quantities of a peptide or protein is often made possible by radiolabeling procedures. The most common labeling strategy, radioiodination with ^{125}I, is relatively straightforward to perform and low in cost. Thus, a peptide with two tyrosine residues can be radiolabeled with ^{125}I to a specific radioactivity of 5×10^9 cpm/μg (see Support Protocol 2), allowing its detection at subfemtomole levels, or several orders of magnitude lower than absorbance- or fluorescence-based detection methods (*UNIT 3.1*).

The main drawback of radioiodination (other than the need to work safely with a potentially hazardous radioisotope) is that some peptides may be unreactive or may be damaged by the introduction of iodine atoms (each as large as a phenyl ring) or by the oxidative conditions typically employed during the labeling reaction. These problems may be more acute for a large protein with native tertiary structure and more oxidation-sensitive residues (tryptophan, methionine, and cysteine) than for a short peptide; nevertheless, even small peptides can be significantly altered in their biological activities by iodination. In most cases these problems can be obviated by modifying the radiolabeling strategy (see Table 3.3.1).

Standard radioiodination methods rely on the oxidation of ^{125}I (in the form of $Na^{125}I$) to its reactive state (which can be considered "I^+") for electrophilic addition to the phenolic and imidazole side chains of the amino acids tyrosine and histidine, respectively (see Fig. 3.3.1). At a pH between 6 and 7, tyrosine labeling predominates, whereas at a higher pH there is also labeling on histidine (Wolff and Covelli, 1969; Tsomides and Eisen, 1993). Different methods of oxidizing the ^{125}I include Iodo-Beads (a convenient immobilized form of chloramine T; Basic Protocol 1), soluble chloramine T, and Iodogen (a water-insoluble reagent plated onto tubes prior to use; Alternate Protocol 1). A gentler oxidative procedure uses the enzyme lactoperoxidase along with hydrogen peroxide (Alternate Protocol 2). All four methods can be expected to give reasonable specific radioactivities (lowest for lactoperoxidase), and their relative merits have been assessed for some peptides (Thean, 1990; Rizzino and Kazakoff, 1991).

For peptides that lack tyrosine and histidine or that are unacceptably altered by standard iodination procedures, an alternate strategy is to label free amino groups (lysine side chains and the peptide N-terminus, unless it is proline or blocked) using an iodinated acylating agent (e.g., Bolton-Hunter reagent; see Basic Protocol 2 and Fig. 3.3.1; Bolton and Hunter, 1973). This mild coupling reaction avoids direct contact of the peptide or protein with harsh oxidizing conditions, but the resulting labeled product is likely to have a lower specific radioactivity (because Bolton-Hunter reagent suffers competing hydrolysis during the reaction); a second drawback is the relatively high cost of this reagent. A similar though less commonly tried modification using an iodinated amidinating reagent (Wood et al., 1975) has the potential advantage that the positive charge on lysine (pK_a~10) is retained in the final labeled product (an amidine with pK_a 11.5 to 12.5).

Besides labeling with ^{125}I, it is straightforward to radiolabel peptides or proteins with ^{14}C or ^3H using various methods (see Table 3.3.1). This may be desirable if all iodination options fail, or if it is critical to obtain a radioactive product devoid of structural alterations. Usually the products of ^{14}C or ^3H labeling will have much lower specific radioactivities than the products of ^{125}I labeling, reflecting the different disintegration rates for these radioisotopes (~35,000 and 75 times lower for ^{14}C and ^3H, respectively, than for ^{125}I). Also, counting β emissions requires liquid scintillants and is therefore less convenient than counting γ emissions, which requires no additives to the sample. However, the longer half-lives of ^{14}C and ^3H (5760 and 12.26 years, respectively) render the resulting products useful for much longer than ^{125}I-labeled products ($t_{1/2}$ ~60 days).

A selection of protocols for labeling a peptide with ^{14}C or ^3H is given in this unit: acetylation on amino groups with acetic anhydride (see Basic Protocol 3), reductive alkylation on amino groups with formaldehyde and sodium cyanoborohydride (see Alternate Protocol 3), and alkylation on cysteine with iodoacetic acid or iodoacetamide (see Alternate Protocol 4). Each of these procedures results in modification of the peptide structure, albeit on different residues than does standard radioiodination (Fig. 3.3.1). To label a peptide without altering its structure, two alternatives are available: ^3H labeling by catalytic reduction of (nonradioactive) ^{127}I-labeled tyrosine (see Basic Protocol 4), and de novo peptide synthesis using a radiolabeled (and suitably protected) amino acid

UCU IAS LIBRARY

derivative (see Basic Protocol 5; Stewart and Young, 1984). These methods are much more laborious and costly, but may be useful when modification of a peptide's structure is unacceptable.

Critical Parameters and Troubleshooting

Amount of radioisotope

The amount of radioisotope to be used is often a compromise among the desire for high specific radioactivities (curies per millimole or cpm per microgram), a practical limit to the amount of radioactivity that can be afforded and safely handled, and (particularly for a large protein) any risk of altering tertiary structure by excessive substitution or radiation damage. Commonly, 1 to 5 mCi of ^{125}I (or ^{131}I) is used; however, it is possible to scale up the ^{125}I and increase specific radioactivities proportionately, because iodide is almost always limiting relative to iodinatable sites (e.g., 2 mCi ^{125}I is only ~1.2 nmol). Conversely, lowering the amount of peptide or protein will also improve specific radioactivities, but lowering the amount of substrate too far will cause unacceptable losses of material by nonspecific adsorption to tubes; a practical lower limit is 5 to 10 μg of peptide.

Sensitive amino acid residues

During oxidative iodinations, the amino acid residues at greatest risk for damage are tryptophan, methionine, and cysteine. The problem is greatest for tryptophan, as an oxidized indole ring cannot be restored by subsequent treatment with a reducing agent such as dithiothreitol. When these residues are critical for peptide function, the lactoperoxidase or Bolton-Hunter methods may be preferred for ^{125}I labeling.

Iodination work-up

Several options are available for the removal of unbound radioactive iodide from a peptide once the reaction is complete, including solid-phase extraction (e.g., using Sep-Pak cartridges; Basic Protocol 1), anion-exchange chromatography (*UNIT 8.2*), gel filtration (*UNIT 8.3*), and dialysis (*APPENDIX 3B*). The choice of desalting procedure will depend on the peptide, particularly its size and charge. For small peptides the Sep-Pak method works well, because the peptide binds to the Sep-Pak cartridge, allowing extensive washing to remove all traces of free iodide (Tsomides and Eisen, 1993). For larger peptides and proteins, anion-exchange chromatography (e.g., Dowex column) and gel filtration are popular, as the peptide is eluted immediately while iodide is retained on the column. To ensure that the removal of unbound iodide is complete, many investigators precipitate a radiolabeled protein with trichloroacetic acid (TCA) to determine the proportion of TCA-precipitable counts. Dialysis is a time-consuming but effective method for removing traces of iodide, provided the peptide is large enough to be retained by the dialysis membrane used.

Note that for highly negatively charged peptides, there may be some retention on anion-exchange columns. For peptides of 10 to 20 amino acids, separation of labeled peptide from free iodide may be poor on an anion-exchange or gel-filtration column because the peptide's molecular mass is so close to the exclusion limit of the column. For such peptides the Sep-Pak method may be advantageous.

Once unbound radioactive iodide has been removed, the labeled peptide mixture can be used as is, or the individual products can be separated from each other as well as from the unreacted peptide that is almost always present (see Support Protocol 2). If unlabeled peptide is removed by reversed-phase HPLC, the specific radioactivity of the labeled peptide will be maximized and will equal the specific radioactivity of carrier-free ^{125}I (~2000 Ci/mmol) multiplied by the number of ^{125}I atoms incorporated per peptide molecule.

Poor iodination results

Table 3.3.2 gives potential causes and possible remedies for poor iodination results.

Anticipated Results

In radioiodinating a purified protein, it is customary to aim for incorporation of up to one ^{125}I atom per protein molecule, i.e., ~20 μCi/μg for a protein of molecular mass 100,000 Da. For short peptides it is rare to label all the molecules, but specific radioactivities are typically on the same order (e.g., 20 μCi/μg for a peptide of mass 1000 Da, representing only 1% of the peptide labeled and 99% unlabeled). More labeled peptide is obtained when nonradioactive ^{127}I is added during the iodination reaction (see Support Protocol 1). For example, Figure 3.3.2 shows the results of iodinating an eight-residue peptide with 20 mCi ^{125}I and 5 μmol ^{127}I, followed by separating the products from unlabeled peptide by HPLC.

POLARY CENTRAL UNIVERSITY LIBRARY

Table 3.3.2 Troubleshooting the Iodination Reaction

Problem	Potential cause	Possible remedy
Poor iodine incorporation	Peptide contains reducing agent, e.g., from synthesis/cleavage	HPLC-purify or desalt peptide prior to use
	Peptide contains detergent or other component that becomes iodinated	HPLC-purify peptide
	Tyrosine/histidine residues are unreactive because of tertiary structure	Treat peptide with denaturing agent Use alternate strategy, e.g., Bolton-Hunter reagent
	Iodo-Beads are unreactive	Test for color (see Support Protocol 1); use fresh lot
Poor peptide recovery	Solubility problem	Change solvent system or adjust pH
	Failed elution from Sep-Pak	Replace with ion exchange, gel filtration, or dialysis
	Nonspecific losses	Increase amount of peptide or add carrier protein (after iodination reaction)
		Use chloramine T instead of Iodo-Beads
Labeled peptide unreactive	Tyrosine is crucial for activity	Iodinate for shorter time and/or use less label
		Use alternate strategy, e.g., Bolton-Hunter reagent
	Other peptide damage	Use alternate strategy

Time Considerations

Each of the listed protocols is easily completed within a few hours (not including HPLC purification if chosen), and with practice can be done in even less time. The resulting radioactive peptides will, of course, decay according to the half-life of the radioisotope used (or in some cases faster, if there is radiation damage). To calculate specific radioactivity after t days, simply multiply the initial specific radioactivity by $2^{(-t/t_{½})}$, where $t_{½}$ is the half-life of the radioisotope (60 days for ^{125}I).

Literature Cited

Bolton, A.E. and Hunter, W.M. 1973. The labelling of proteins to high specific radioactivities by conjugation to a ^{125}I-containing acylating agent. *Biochem. J.* 133:529-539.

Chersi, A., Trinca, M.L., and Camera, M. 1988. ^{14}C-labeling of synthetic peptides. *J. Immunol. Methods* 110:271-273.

Crestfield, A.M., Moore, S., and Stein, W.H. 1963. The preparation and enzymatic hydrolysis of reduced and S-carboxymethylated proteins. *J. Biol. Chem.* 238:622-627.

Edelhoch, H. 1962. The properties of thyroglobulin. VIII. The iodination of thyroglobulin. *J. Biol. Chem.* 237:2778-2787.

Eisen, H.N. and Keston, A.S. 1949. The immunologic reactivity of bovine serum albumin labeled with trace amounts of radioactive iodine (^{131}I). *J. Immunol.* 63:71-80.

Fraenkel-Conrat, H. and Colloms, M. 1967. Reactivity of tobacco mosaic virus and its protein toward acetic anhydride. *Biochemistry* 6:2740-2745.

Fraker, P.J. and Speck, J.C., Jr. 1978. Protein and cell membrane iodinations with a sparingly soluble chloroamide, 1,3,4,6-tetrachloro-3α,6α-diphenylglycoluril. *Biochem. Biophys. Res. Commun.* 80:849-857.

Gahmberg, C.G. and Andersson, L.C. 1977. Selective radioactive labeling of cell surface sialoglycoproteins by periodate-tritiated borohydride. *J. Biol. Chem.* 252:5888-5894.

Greenwood, F.C., Hunter, W.M., and Glover, J.S. 1963. The preparation of ^{131}I-labelled human growth hormone of high specific radioactivity. *Biochem. J.* 89:114-123.

GGU IASL CENTRAL UNIVERSITY LIBRARY

Hubbard, A.L. and Cohn, Z.A. 1975. Externally disposed plasma membrane proteins. I. Enzymatic iodination of mouse L cells. *J. Cell Biol. 64:438-460.*

Hunter, W.M. and Greenwood, F.C. 1962. Preparation of iodine-131 labelled human growth hormone of high specific activity. *Nature* 194:495-496.

Jentoft, N. and Dearborn, D.G. 1983. Protein labeling by reductive alkylation. *Methods Enzymol.* 91:570-579.

Marchalonis, J.J. 1969. An enzymatic method for the trace iodination of immunoglobulins and other proteins. *Biochem. J.* 113:299-305.

Marchalonis, J.J., Cone, R.E., and Santer, V. 1971. Enzymic iodination: A probe for accessible surface proteins of normal and neoplastic lymphocytes. *Biochem. J.* 124:921-927.

Markwell, M.A.K. 1982. A new solid-state reagent to iodinate proteins: Conditions for the efficient labeling of antiserum. *Anal. Biochem.* 125:427-432.

Rizzino, A. and Kazakoff, P. 1991. Iodination of peptide growth factors: Platelet-derived growth factor and fibroblast growth factor. *Methods Enzymol.* 198:467-479.

Sarin, V.K., Kent, S.B.H., Tam, J.P., and Merrifield, R.B. 1981. Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction. *Anal. Biochem.* 117:147-157.

Stewart, J.M. and Young, J.D. 1984. Solid Phase Peptide Synthesis. Pierce Chemical Co., Rockville, Ill.

Tack, B.F. and Wilder, R.L. 1981. Tritiation of proteins to high specific activity: Application to radioimmunoassay. *Methods Enzymol.* 73:138-147.

Thean, E.T. 1990. Comparison of specific radioactivities of human α-lactalbumin iodinated by three different methods. *Anal. Biochem.* 188:330-334.

Tsomides, T.J. and Eisen, H.N. 1993. Stoichiometric labeling of peptides by iodination on tyrosyl or histidyl residues. *Anal. Biochem.* 210:129-135.

Wood, F.T., Wu, M.M., and Gerhart, J.C. 1975. The radioactive labeling of proteins with an iodinated amidination reagent. *Anal. Biochem.* 69:339-349.

Wolff, J. and Covelli, I. 1969. Factors in the iodination of histidine in proteins. *Eur. J. Biochem.* 9:371-377.

Key References

Means, G.E. and Feeney, R.E. 1971. Chemical Modification of Proteins. Holden-Day, Inc., San Francisco, Calif.

Excellent overview of side chain chemistries, including but not limited to radiolabeling procedures.

Tsomides and Eisen, 1993. See above.

Describes use of Sep-Pak to remove free iodine, HPLC to separate iodinated products, Edman degradation to identify products, and pH dependence of tyrosine vs. histidine iodination using Iodo-Beads.

Contributed by Ton N.M. Schumacher and Theodore J. Tsomides
Massachusetts Institute of Technology
Cambridge, Massachusetts

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Assays for Total Protein

Total protein assays are used to analyze hundreds of industrial, agricultural, and biotechnology products. They are also basic for research purposes, especially for determining the specific activity (i.e., total activity/total protein) of enzymes, antibodies, and lectins. Clearly, accuracy and precision of specific activity measurements depend as much on accurate measurement of total protein as on determination of total activity.

In performing total protein assays, there are five issues of concern: (1) sensitivity and technique of the method, (2) clear definition of units, (3) interfering compounds, (4) removal of interfering substances before assaying samples, and (5) correlation of information from various techniques. These concerns are discussed at length in the Commentary of this unit.

This unit describes three copper-based assays to quantitate total protein: the biuret method (Basic Protocol 1), a variation of the Lowry method (Hartree-Lowry method; Basic Protocol 2), and the bicinchinonic acid (BCA) assay (Basic Protocol 3). Acid hydrolysis of a protein is coupled with ninhydrin detection to quantitate amino acid content of a sample (Basic Protocol 4). Ultraviolet spectrophotometry is used to measure total protein (Basic Protocol 5) and evaluate samples for the presence of contaminants. The Coomassie dye binding, or Bradford, assay (Basic Protocol 6) is a quite simple assay and frequently is quite sensitive, although it sometimes gives a variable response depending on how well or how poorly the protein binds the dye in acid pH. Finally, dry weight measurement (Basic Protocol 7) is used to quantitate pure protein. Support protocols describe a technique for heat sealing glass tubes (Support Protocol 1) for acid hydrolysis, sample dialysis in polyacrylamide gel wells to remove low-molecular-weight contaminants (Support Protocol 2), and TCA precipitation to precipitate and concentrate proteins and remove low-molecular-weight contaminants (Support Protocol 3).

BIURET ASSAY FOR QUANTITATION OF TOTAL PROTEIN

BASIC PROTOCOL 1

Biuret assays are the oldest and least sensitive assays for total protein among the colorimetric (spectrophotometric) methods. Nevertheless, the biuret assay is still frequently used because it is easy to perform, reagents are easily prepared, and the assay is markedly less susceptible to chemical interference than other copper-based assays. The assay is based on polypeptide chelation of cupric ion (colored chelate) in strong alkali. Catalog listings for many commercially available proteins and enzymes include their specific activities based on biuret protein (mg-biuret-protein^{-1}). In general, biuret assays are useful for samples containing ~1 to 10 mg protein/ml, which is diluted ~5-fold by the added reagent to give a concentration of 0.2 to 2 mg/ml final assay volume (f.a.v.). Most proteins produce a deep purple color, with maximum absorbance (λ_{max}) at 550 nm.

Materials

Calibration standard: e.g., 10 to 20 mg/ml BSA (see Table 3.4.1)
Buffer or solvent used to prepare the protein-containing sample
Sample containing protein at 1 to 10 mg/ml
Biuret reagent (see recipe)
Spectrophotometer and 1-cm cuvettes

1. Prepare a dilution series of calibration standard in the buffer or solvent used to prepare the sample.

   Use bovine serum albumin (crystallized or lyophilized or one of the Cohn Fraction V preparations which are 96% to 98% protein and 3% to 4% water; Sigma) for the

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON / CENTRAL UNIVERSITY LIBRARY

Table 3.4.1 Slopes of Calibration Plots for the Biuret Assay[a]

Protein	Assay conditions	Slope [A_{650} (μg/ml f.a.v.)$^{-1}$cm^{-1}]
BSA	—	2.3×10^{-4}
BSA	—	2.6×10^{-4}
BSA	Glucose concentration 0.4 to 400 mM	2.7×10^{-4}
BSA	—	2.7×10^{-4}
β-lactoglobulin	—	2.5×10^{-4}
Cellulase enzymes	—	2.2×10^{-4}
BSA[b]	—	3.2×10^{-4}
BSA[c]	—	1.9 to 2.8×10^{-4}
BSA[d]	—	2.3×10^{-4}

[a]Abbreviations: BSA, bovine serum albumin; f.a.v., final assay volume.
[b]Watters (1978).
[c]Goshev and Nedkov (1979).
[d]Beyer (1983).

calibration standard. Concentrations of albumin in the final assay volume (f.a.v.) of 5 ml may range from ~0.40 to 2.50 mg protein/ml—to produce A_{550} readings from ~0.10 to 0.70 in 1-cm. cuvettes—so the dilution series should run from 2 to 12.5 mg/ml.

2. In separate test tubes, add 1 ml of protein-containing sample, of each dilution of the calibration standard, or of the buffer or solvent used to prepare the sample (reference standard) to 4 ml biuret reagent. Incubate 20 min at room temperature.

   The modest amounts of detergent (e.g., deoxycholate or SDS) used to help solubilize proteins from tissue or insoluble biomaterials that will be added with the protein sample have negligible effects on this assay (see references in Table 3.4.1), although they sometimes have marked effects on the Hartree-Lowry assay (Basic Protocol 2). Mild reducing agents and strong oxidizers adversely affect the assay.

3. Measure the net absorbance of the sample, calibration standards, and reference standard at 550 nm (A_{550}) in 1-ml cuvettes. If the spectrophotometer does not automatically give net absorbance readings, subtract the value of the reference standard from those obtained for the sample and calibration standards.

   Sample should not be diluted after the addition of biuret reagent. In general, reaction mixtures for spectrophotometric assays should not be diluted after development even if the color (absorbance) is too intense. Dilution will decrease the required excess of cupric ion, thus upsetting chemical equilibria. If the sample has an A_{550} >1 or 2 (depending on the spectrophotometer), dilute a sample of the original protein and repeat the assay.

4. Prepare a calibration plot by graphing the net A_{550} values for the standards versus protein concentration (μg protein/ml f.a.v.). Determine the protein concentration of the sample by interpolation from the plot.

   Plot the data as shown in Figure 3.4.1 and calculate the slope and its units. (The plots in this unit show net absorbance as the ordinate.) The slope of this plot is directly proportional to the apparent spectrophotometric absorption coefficient (the sensitivity of the assay) in the same units.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

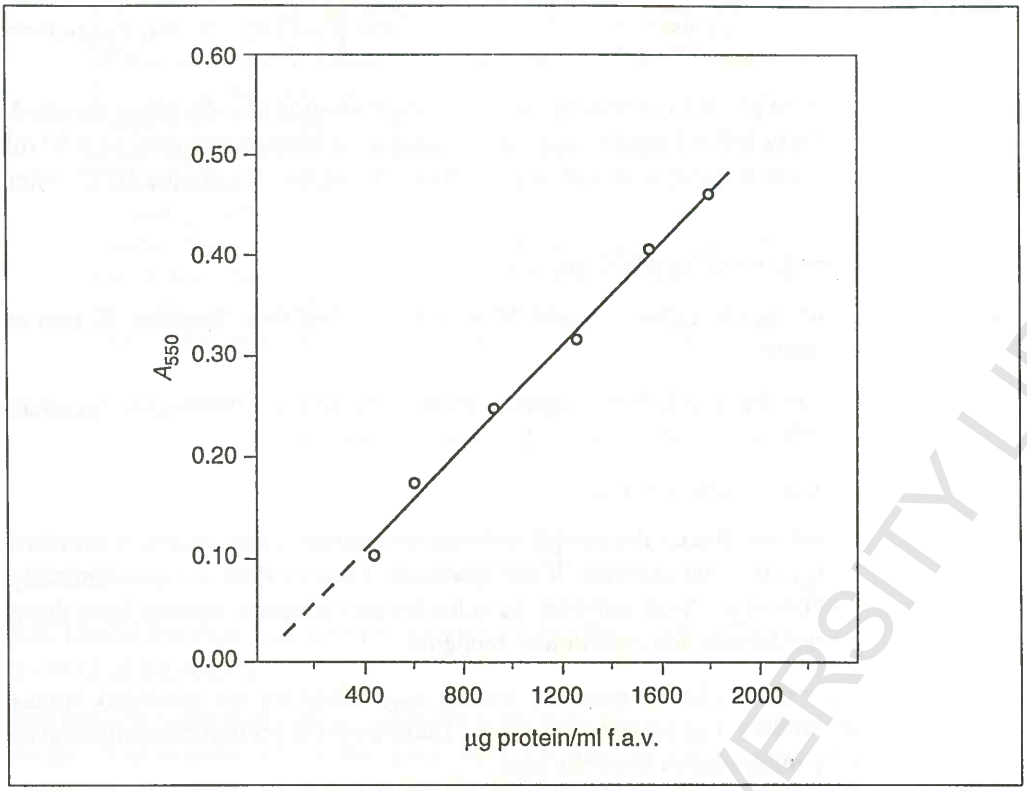


Figure 3.4.1 Biuret assay calibration plot using bovine serum albumin. Slope = 2.26×10^{-4} A_{550} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}.

HARTREE-LOWRY ASSAY FOR QUANTITATION OF TOTAL PROTEIN

BASIC PROTOCOL 2

The original Lowry method for total protein analysis was first described in one of the most cited papers in biochemistry (Lowry et al., 1951). The assay is a colorimetric assay based on cupric ions and Folin-Ciocalteau reagent for phenolic groups. The assay has been reinvestigated many times and sometimes improved. Most of these studies were designed to discern how interfering compounds distort the assay and how detergents solubilize otherwise insoluble proteins. The literature related to this assay was comprehensively reviewed by Peterson (1983). This protocol describes Hartree's version of the Lowry assay (Hartree, 1972). This version uses three reagents instead of five, produces more intense color (increased sensitivity) with some proteins, maintains a linear response over a larger concentration range (30% to 40% greater), is less easy to overload, and overcomes the reagent salt coprecipitation encounter with Lowry reagents. Finally, the reagents formulation offers some advantages in storage stability. It is somewhat less laborious than the original Lowry assay, and it maintains the sensitivity of the original.

Materials

Calibration standard: 300 µg crystalline BSA per ml in water
Buffer or solvent used to prepare the protein-containing sample
Sample containing protein at 100 to 600 µg/ml
Hartree-Lowry reagents A, B, and C (see recipes)

50°C water bath
Spectrophotometer and 1-cm cuvettes

1. Prepare a dilution series of calibration standard in the same buffer or solvent used to prepare the sample to give concentrations of 30 to 150 µg/ml.

BOGAZICI UNIVERSITY LIBRARY

Concentrations of albumin in the final assay volume (f.a.v.) of 5 ml may range from ~30 to 300 µg protein/ 5 ml f.a.v. to produce A_{650} readings from ~0.20 to 0.80.

2. Add 1.0 ml of the protein-containing sample, of each dilution of calibration standard, or of the buffer or solvent used to prepare the sample (reference standard) to 0.90 ml of Hartree-Lowry reagent A in separate test tubes. Incubate 10 min in a 50°C water bath.

3. Cool the tubes to room temperature.

4. Add 0.1 ml of Hartree-Lowry reagent B to each tube and mix. Incubate 10 min at room temperature.

5. Rapidly add 3 ml Hartree-Lowry reagent C to each tube and mix thoroughly. Incubate 10 min in a 50°C water bath, then cool to room temperature.

   The final assay volume is 5.0 ml.

6. Measure the net absorbance the sample, calibration standards, and reference standard at 650 nm (A_{650}) in 1-cm cuvettes. If the spectrophotometer does not automatically give net absorbance readings, subtract the value for the reference solution from those obtained for the sample and calibration standards.

7. Prepare a calibration plot by graphing the net A_{650} values for the standards versus protein concentration (µg protein/5 ml f.a.v.). Determine the protein concentration of the sample by interpolation from the plot.

   Plot the data as shown in Figure 3.4.2 and calculate the slope and its units. The slope of a calibration plot is proportional to the spectrophotometric absorption coefficient, and is a measure of the sensitivity of assay over the range indicated on the horizontal axis. Comparisons can be made with other measures, such as the weight absorption

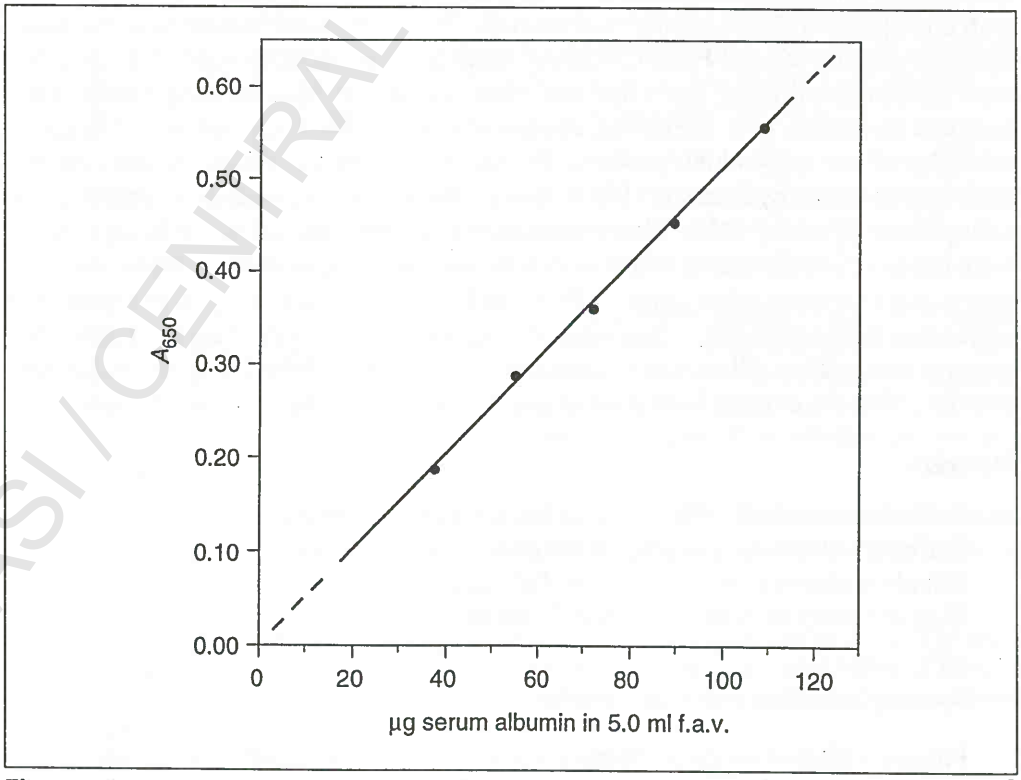


Figure 3.4.2 Hartree-Lowry assay calibration plot using bovine serum albumin. Slope = 2.5×10^{-2} A_{650} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

coefficient ($E^{1\%}$) at the same wavelength. Units in the concentration factor in the denominator of $E^{1\%}$ are (dry g)/100 ml of final assay volume.

The units in this protocol, µg protein/ml final assay volume (f.a.v.), are used to permit comparison with values in the literature. It is not always clear whether the milligram or microgram amounts of protein indicated on the axes of these plots refer to the concentration in the input sample or to the final concentration after the addition of reagent(s). Differences in volumes of input versus final developed samples frequently involve dilution factors of 10^1 to 10^2. As an example, the $E^{1\%}$ values in the original report describing this procedure (see Table 4 in Lowry et al., 1951) are ~230 at 750 nm with 3-mm pathlengths. These values can be converted to units presented in this protocol by reconciling the two means for expressing the slopes (see Equation 3.4.1).

$$E^{1\%}_{1\ cm} = \frac{230\,A_\lambda}{\frac{g}{100\ ml\ f.\ a.\ v.} \times \frac{10^6\ \mu g}{g} \times cm} = \frac{2.3 \times 10^{-2} A_\lambda \times cm^{-1}}{\mu g/ml\ f.\ a.\ v.}$$

Equation 3.4.1

BICINCHINONIC ACID (BCA) ASSAY FOR QUANTITATION OF TOTAL PROTEIN

BASIC PROTOCOL 3

The bicinchinonic acid (BCA) assay for total protein is a spectophotometric assay based on the alkaline reduction of the cupric ion to the cuprous ion by the protein, followed by chelation and color development by the BCA reagent. Either a micro or a semimicro procedure, the latter generating a final assay volume of 2 or 3 ml, may be used; the semimicro procedure is most frequently used.

The BCA assay for total protein is somewhat variable: it has differing sensitivities in response to incubation time, incubation temperature, standard protein used for calibration, and other factors (Smith et al., 1985). Certain classes of compounds such as reducing sugars and ammonium ions interfere with the assay, sometimes severely. However, if interfering compounds are eliminated (e.g., by dialysis; see Support Protocol 2), the BCA assay has a good combination of sensitivity and simplicity, and it has some advantages over the Lowry technique (see Background Information).

Materials

Calibration standard: 1 mg BSA/ml
Buffer or solvent used to prepare the protein-containing sample
Sample containing protein
BCA reagent A/reagent B mix (see recipe)
Spectrophotometer and cuvettes

1. Prepare a dilution series of calibration standard in the buffer or solvent used to prepare the sample to cover the range 0.2 to 1.0 mg/ml.

2a. *For a 2.1-ml final assay volume:* Mix 100 µl of protein-containing sample, of calibrating standard, or of buffer used to prepare the sample (reference standard) with 2 ml BCA reagent A/reagent B mix in separate tubes.

2b. *For a 4.2-ml final assay volume:* Mix 200 µl of protein-containing sample, of calibrating standard, or of buffer used to prepare the sample (reference standard) with 4 ml BCA reagent A/reagent B mix in separate test tubes.

The final assay volume depends on the size of the available cuvettes.

BOU INST / BEIJING UNIVERSITY LIBRARY

3. Incubate the samples and standards 30 min at 37°C, then cool to room temperature.

4. Measure the absorbance of the sample, calibration standards, and reference standard at 562 nm (A_{562}). If the spectrophotometer does not automatically give net absorbance readings, subtract the values for the reference standard from those obtained for for the sample and calibration standards.

5. Prepare a calibration plot by graphing the net A_{650} values for the standards versus protein concentration (µg protein/ml f.a.v; see Fig. 3.4.3). Determine the protein concentration of the sample by interpolation from the plot.

Figure 3.4.3 shows a calibration plot prepared using BSA as the protein standard and a 2.1-ml final assay volume (f.a.v.). The slope of the plot is a measure of the sensitivity of the assay. The slopes of plots such as Figure 3.4.3 are also related to the absorption coefficients of the colored complexes. The units attached to the slope must take into account the extent of dilution and be consistent with the protocol used to develop the colored complexes. The value of the slope can be used to compare the different methods—e.g., Smith et al. (1985) compared the BCA and Lowry methods using BSA as calibrating protein. For the BCA method, the slope was 1.5×10^{-2} to 2.6×10^{-2} A_{562} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}; for the Lowry method, the slope was 1.6 to 2.3×10^{-2} A_{750} µg protein/ml f.a.v.)$^{-1}$ cm^{-1}.

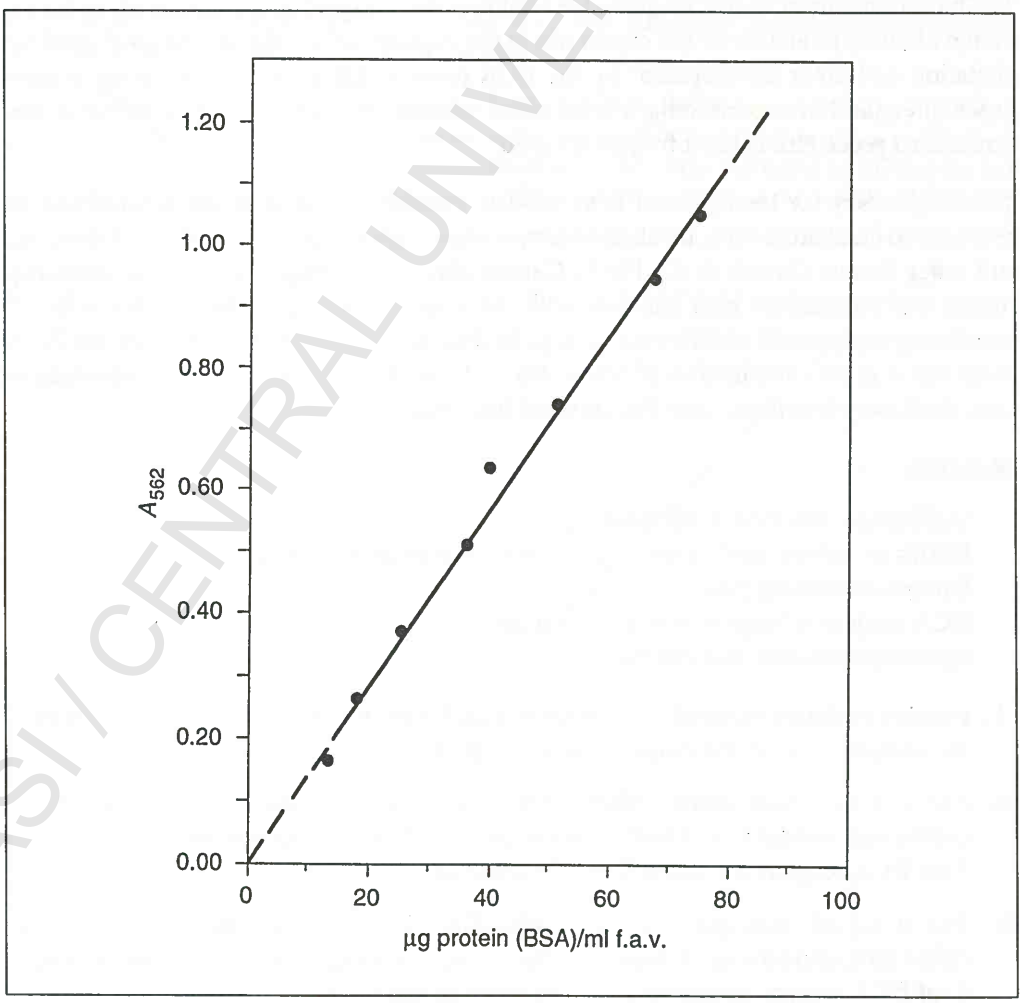


Figure 3.4.3 Bicinchinonic acid assay calibration plot using bovine serum albumin. Slope = 1.5×10^{-2} A_{562} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

ACID DIGESTION–NINHYDRIN METHOD FOR QUANTITATION OF TOTAL PROTEIN

In this assay, proteins are hydrolyzed to amino acids with 6% sulfuric acid at 100°C. The hydrolysates are neutralized and the amino acids quantitated by ninhydrin derivatization and spectrophotometry at 570 nm (A_{570}). This method is resistant to interference by phenolic and related compounds that may affect Folin reaction–dependent analyses.

Individual amino acids do not produce equal "color yields" upon reaction with ninhydrin. However, most proteins—except those with unusual compositions such as high hydroxyproline and proline (collagens), extraordinary sulfur contents (cysteine; e.g., alcohol dehydrogenase and thionein), or densely glycosylated proteins (e.g., invertase and glycophorin)—give reasonable results using leucine as a calibration standard. Ammonium ion generates strong color with ninhydrin, nearly equivalent to that produced by leucine. However, TCA precipitation of the proteins prior to acid hydrolysis will remove NH_4^+, which will be left in the supernate (see Support Protocol 3).

Materials

H_2SO_4
NaOH
Sample containing protein
Calibration standard: 0.2 mg/ml leucine, diluted ½ to 1/20
Ninhydrin reagent (see recipe)
Isopropanol diluent: 50% (v/v) isopropanol in water

Heat-sealable tubes
100°C water bath

1. Based on the percentage H_2SO_4 (e.g., 6%) to be used in this assay, determine the volume of NaOH required to neutralize the reaction.

   For example, 6% H_2SO_4 is ~2.3 N (H^+), so it would take ~15.3 ml of 6% (w/v) NaOH (1.5 N) to completely neutralize 10 ml of 6% H_2SO_4.

   In this protocol, % H_2SO_4 refers to a v/v dilution of concentrated H_2SO_4.

   The total volume of the reaction mixture after the addition of acid, base, and ninhydrin reagent to a protein solution containing 10 to 20 mg protein should be close to 3 but no more than 6 ml. If necessary, use more concentrated acid and base solutions to prevent excessive dilution of the sample.

2. In a heat-sealable tube, add H_2SO_4 to an 0.02- to 0.05-mg protein sample to give a final concentration of 3% acid. Seal the tube under a nitrogen blanket (see Support Protocol 1) and incubate 12 to 15 hr (overnight) at 100° to 105°C.

   Protein in aqueous solution may be used directly if it does not contain many additional amino acids and oligopeptides. If the sample is dilute or contains contaminants or interfering compounds, the protein can be precipitated and concentrated using TCA (see Support Protocol 3). After the supernatant is removed, 3% H_2SO_4 can be added directly to the TCA precipitate for hydrolysis.

3. Open the tube and add the appropriate volume of 6% NaOH, as determined in step 1, to neutralize the H_2SO_4 used in hydrolysis.

4. Add 2 ml ninhydrin reagent to the neutralized sample, and to equal volumes of diluted calibration standards, and of the buffer (reference standard) in separate test tubes. Incubate the mixture 20 min in 100°C (boiling) water bath. Cool to room temperature.

   Calibration and reference standards should be light blue, samples a darker purplish blue. Very dark samples should be diluted by a factor of 2 to 5 with isopropanol diluent

BOSTON COLLEGE UNIVERSITY LIBRARY

(dilute all standards to the same extent). Dilution may be performed after incubation with the reagent, and it is not necessary to repeat the assay in such cases.

6. Measure the net absorbance of sample and calibration and reference standards at 570 nm (A_{570}). If the spectrophotometer does not automatically give net absorbance readings, subtract the value for the reference standard from those obtained for sample and calibration standards.

   The absorbance of a sample dominated by hydroxyproline and proline may be read at 440 nm; a separate set of calibration standards should be prepared with these amino acids.

7. Prepare a calibration plot by plotting the adjusted values versus protein concentration (µg protein/ml f.a.v.). Determine the protein concentration of the sample by interpolation from the curve.

   The slope of ninhydrin color yield (A_{570}) with most amino acids is 20 to 25 A_{570} µmol amino acid/ml f.a.v.$^{-1}$ cm^{-1}. As a general guide, use 0.02 to 0.05 mg protein per assay.

SUPPORT PROTOCOL 1

HEAT SEALING GLASS TUBES

For acid hydrolysis at 100° to 104°C, air (oxygen) should be excluded and evaporation should be prevented by carrying out the hydrolysis in sealed tubes. Ordinary glass test tubes, ~10 to 12 mm diameter and 10 cm long, can be used for this purpose. Figure 3.4.4 summarizes the procedure.

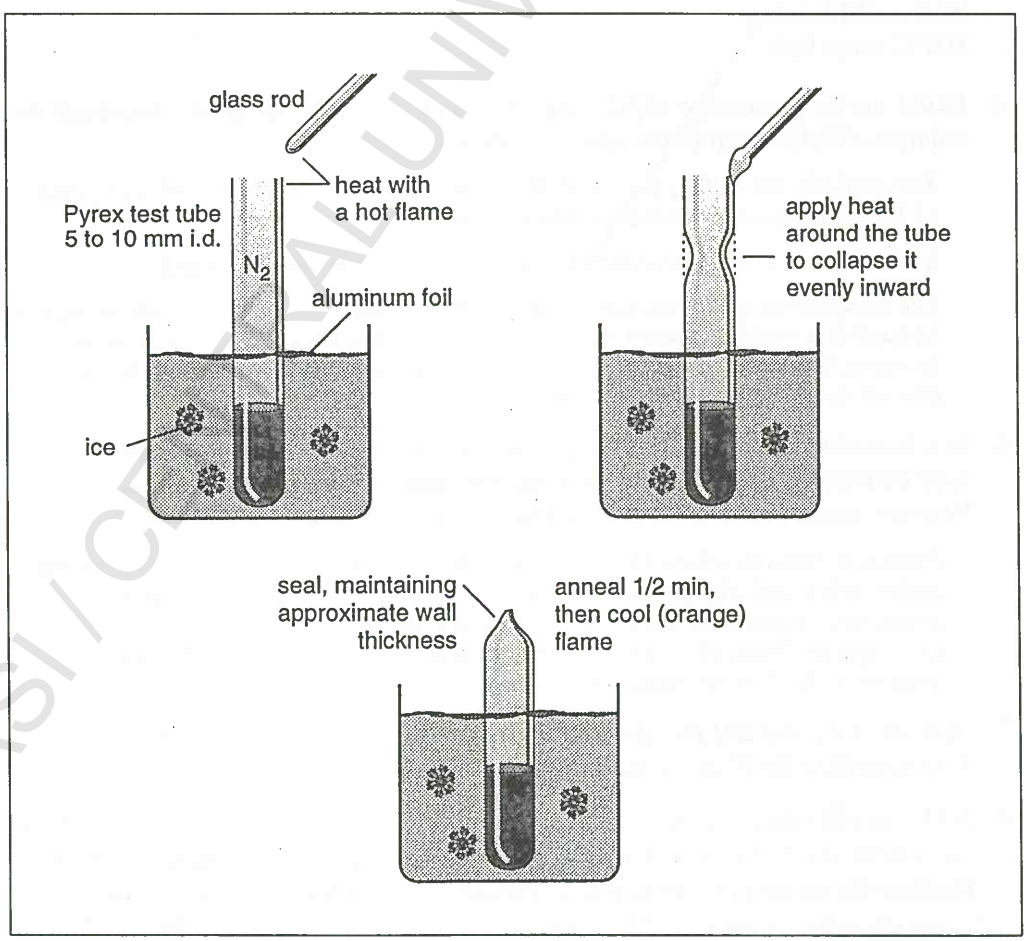


Figure 3.4.4 Heat sealing test tubes for hydrolysis.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Materials

Pyrex glass test tube, 10 to 12 mm diameter × 10 cm long
Nitrogen gas
Glass rod
Gas-oxygen flame source
Triangular file

1. Set a Pyrex test tube containing sample to be hydrolyzed in ice and briefly blanket the tube with nitrogen gas.

   CAUTION: *Adjust the force of the nitrogen stream carefully to avoid splashing the acid solution out of the tube or up the sides of the tube.*

2. Using a gas-oxygen flame, fuse a small glass rod to the lip of the test tube to act as a handle.
3. Use the flame to apply heat around the body of the tube, ~2 cm below the lip until the tube softens and collapses.
4. Slowly withdraw the "handle" and the scrap upper end, maintaining test tube wall thickness around the seal.
5. When the seal is fused, reduce the flame temperature down to an orange, then a red flame and continue to flame the seal for ~½ min.

   This should anneal the glass. Do not build a thick seal, which is likely to crack, or draw out the seal so that it becomes thin and likely to blow out.

6. After acid hydrolysis, break the tubes open as follows. First make a deep scratch with a sharp triangular file, about one-third of the way around the tube. Then heat a glass rod until it is white hot, wet the scratch, and apply the white-hot glass rod to the scratch.

   This quite reliably cracks the tube all round and the top easily lifts off.

ULTRAVIOLET ABSORPTION TO MEASURE TOTAL PROTEIN

BASIC PROTOCOL 5

Proteins contain tyrosine and tryptophan side chains that are fairly strong absorbers of light in the 275- to 280-nm (ultraviolet) region. Consequently, after suitable dilution to produce on-scale absorbance readings, total protein can be estimated from UV absorbance spectra using quartz or fused silica cuvettes. Phenylalanine, which has an aromatic side chain, is only weakly absorbing and is usually neglected for most purposes. Weight absorption coefficients, $E^{1\%}$ values (concentration of protein in dry grams per 100 ml volume or weight/volume percent) range between ~3 and 30 A_{280} units $(gm/100\ ml)^{-1}\ cm^{-1}$ for most proteins. A useful survey of $E^{1\%}$ values for individual proteins at 275 to 280 nm is available in Sober (1970). If the molecular weight of proteins is known, the molar absorption coefficient, ε_λ in moles/liter is related to the $E^{1\%}$ values at the same wavelength, by Equation 3.4.2.

$$10\ \varepsilon_\lambda = E_\lambda^{1\%} \times \text{molecular weight}$$

Equation 3.4.2

Molar absorption coefficients can be calculated from amino acid composition based on tyrosine and tryptophan contributions to the total molecule. The $\varepsilon_{275\text{-}280}$ values for tyrosine (in neutral or acidic solution) and for tryptophan are close to 1470 and 5700 absorbance

B.C.U. Int. Z.[illegible] CHILE
BCU IAS CENTRAL UNIVERSITY LIBRARY

units M^{-1} cm^{-1}, respectively. There are minor variations in these values depending on the peptide and the solvent (Bencze and Schmid, 1957).

For example, an 0.04% to 0.12% solution of protein having 4 to 12 tyrosine groups, 2 to 6 tryptophans, and an ε_{280} value of ~8, should have an A_{280} of ~0.2 to 1 with a 1-cm path length cuvette. Hence UV absorption around the 280-nm band provides a fairly sensitive, convenient means for detecting and quantitating pure proteins or mixtures of pure proteins having concentrations in the range of 0.1 to 1 mg/ml.

Materials

Protein sample, pH <8
Fused-silica or quartz cuvettes

1. Turn on the spectrophotometer, set the wavelength to 280 nm, and allow it to warm up 30 min.
2. Dilute the protein sample, if necessary, to give an A_{280} between 0.2 and 1.

   Some spectrophotometers may give precise measurements of A_{280} up to 2 absorption units but in general, conditions and concentrations should be adjusted so the instrument is not forced to read absorbances >0.9 to 1.0.

3. Read the absorbance of sample and reference standard (buffer or solvent) at 280 nm in fused-silica or quartz cuvettes.

   Glass and plastic cuvettes are opaque below the near UV range, ~350 mm.

   A complete UV absorption spectrum, ~230 or 240 nm to >310 nm, may be very helpful for detecting interfering compounds. For a solution of one or more pure proteins at acid pH, there should be a deep trough at 250 nm, a peak around 275 to 280 nm, and very little absorption above 310 nm, unless the proteins contain cofactors or prosthetic groups.

4. If necessary, dialyze or precipitate the sample to remove contaminants and interfering compounds and reread the A_{280}.

BASIC PROTOCOL 6

COOMASSIE DYE–BINDING ASSAY (BRADFORD ASSAY) TO MEASURE TOTAL PROTEIN

Coomassie dye (Brilliant blue G250) binds to protein molecules in acid pH by two means. The triphenylmethane group binds to nonpolar structures in proteins, and the anion sulfonate groups interact with protein cationic side chains (e.g., arginine and lysine side chains) in acid pH (Lovrien et al., 1995). The color change produced when the dye binds to proteins provides a measure of total protein, which is quite sensitive in the case of albumin and certain globular proteins (Bradford, 1976; Sedmak and Grossberg, 1977). The Coomassie dye-binding assay, or Bradford assay, also responds to some interfering substances which are generally unknown unless specifically tested for (Van Kley and Hale, 1977; Kirazov et al., 1993). Nevertheless, because of its apparent simplicity and sensitivity towards many proteins, the Bradford assay is popular and widely used. Using bovine serum albumin (ideally well behaved) to calibrate the Bradford assay produces a calibration plot with a slope of 4.5 to 5.5 × 10^{-2} (μg protein/ml f.a.v.)$^{-1}$cm^{-1}. Hence, it is somewhat more sensitive, by a factor of roughly two or three, than the values generally quoted for Lowry, Hartree-Lowry, or BCA assays (see Fig. 3.4.2 and see Table 3.4.2).

BCU IAS CENTRAL UNIVERSITY LIBRARY

Materials

Calibration standards: 1.5 mg/ml BSA and 1.5 mg/ml lysozyme
Buffer or solvent used to prepare the protein-containing sample
Sample containing protein
Coomassie dye reagent (see recipe) *or* commercial Coomassie reagent (Bio-Rad, Pierce)

1. Prepare a dilution series of calibration standards in the buffer or solvent used to prepare the sample to cover the range 150 to 750 µg protein/ml.

   Depending on the kind of protein being measured, it may be useful to calibrate with various proteins including proteins related to or even a purified preparation of the protein being analyzed. For example, if collagen is the analyte, various collagens, high-glycine-hydroxyproline polymers, should be used as calibrating standards. Bovine serum albumin (BSA) is often used as a calibration standard, but it has greater general dye-binding capacity than most proteins.

2. Add 100 µl of protein-containing sample, calibration standard, or of buffer used to prepare the sample (reference standard) to 5 ml Coomassie dye reagent. Mix. Incubate 10 min at room temperature.

   The ratio of sample to reagent may range from 1:20 to 1:50 (v/v). A micro assay can be performed using 5 to 10 µl sample in 250 to 500 µl Coomassie dye reagent.

   If a commercially prepared Coomassie reagent is used, follow the manufacturer's instructions.

3. Measure the absorbance of the sample, calibration standards, and reference standard at 595 nm (A_{595}). If the spectrophotometer does not automatically give net absorbance readings, subtract the values for the reference standard from those obtained for the sample and calibration standards.

4. Prepare a calibration plot by graphing the net A_{595} values for the standards versus protein concentration (µg protein/ml f.a.v.). Determine the protein concentration of the sample by interpolation from the plot.

   At maximal sensitivity (with BSA as a calibration standard) the Coomassie dye-binding assay produces a calibration plot with a slope (sensitivity) of ~4×10^{-2} A_{595} (µg/ml f.a.v.)$^{-1}$ cm^{-1} and falls to a third or half of that value with less responsive proteins.

DRY WEIGHT DETERMINATION TO MEASURE TOTAL PROTEIN

BASIC PROTOCOL 7

Drying a protein-containing sample in a 104° to 106°C oven for a few hours removes water and volatile materials. The sample is then weighed on a balance to measure the aggregate weight of protein plus whatever nonvolatile material remains, such as salts and many buffers.

Materials

Protein sample
Small weighing bottle or beaker
104° to 106°C oven

1. Dry a small weighing bottle or small beaker by heating it 10 min in a 104° to 106°C oven.

2. Cool the weighing container in a desiccator 10 min.

3. Weigh the container on a balance to the nearest 0.1 mg (tare weight).

BCU IRSI / CENTRAL UNIVERSITY LIBRARY

4. Add a 0.5- to 3-ml sample of protein solution to the container.

   Buffer salts, polysaccharides, the salt forms of amino acids, and high molecular weight pigments and sugars may not be driven off in a 4- to 6-hr drying cycle (see step 5). They should be removed from the sample by dialysis, ion-exchange chromatography, or precipitation before it is dried (see Support Protocols 2 and 3).

5. Dry the sample and container 4 to 6 hr (or overnight) at 104° to 106°C.

   Usually it is not necessary to carry out more than one cycle of drying, cooling, and weighing to approach a constant weight.

   Overnight drying is required for samples to 5 to 10 ml.

6. Cool the container and reweigh it (dried weight).

7. Calculate the dry weight of the sample as: dry sample weight = dried weight of container and protein – tare weight.

   The first dry weights measured after 6 hr usually remain constant, but if volatile salts such as ammonium chloride and ammonium formate are present, the sample may have to be dried longer or for more cycles to completely drive them off.

   Dry weight is reliable to within ~3% for net weights of 2 to 4 mg protein and to within 1% to 2% for ≥5 mg protein.

SUPPORT PROTOCOL 2

GEL DIALYSIS OF PROTEIN SAMPLES

This protocol describes a simple method for dialyzing small interfering molecules and salts away from protein samples using an agarose or polyacrylamide gel (see Fig. 3.4.5; Freifelder and Better, 1982). Somewhat dense, well cross-linked gels are used to keep proteins from penetrating the gel while unwanted salts and low-molecular-weight compounds diffuse (are "dialyzed away") into the gel. This is a micro to semimicro method, an alternative to conventional dialysis (Craig, 1967), and to the use of small pressure-membrane disposable kits or microcentrifuge filtration units with MWCO values of 5,000 to 10,000.

Materials

Polyacrylamide gel (*UNIT 10.1*)
Sample containing protein

Small glass rods or test tubes
Beaker

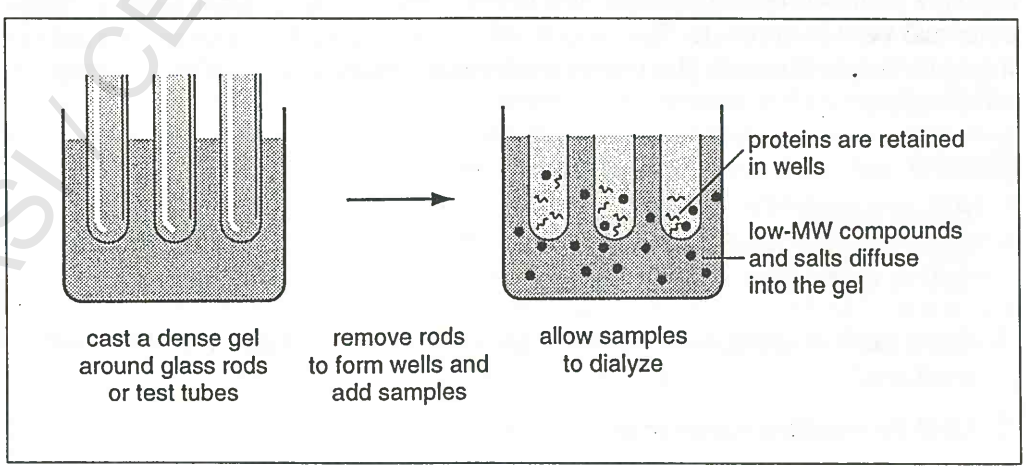

proteins are retained in wells
low-MW compounds and salts diffuse into the gel
cast a dense gel around glass rods or test tubes
remove rods to form wells and add samples
allow samples to dialyze

Figure 3.4.5 Dialysis by out-diffusion of low-molecular-weight compounds into gels cast in small cavities (Beyer, 1983).

BCU IASI CENTRAL UNIVERSITY LIBRARY

1. Suspend small glass rods or test tubes inside a beaker.

   The glass rods or test tubes are used to form wells that will hold the protein sample. Use a size that will create a well of sufficient volume to hold the sample.

2. Prepare a 12% to 15% polyacrylamide gel cross-linked with 0.8% to 1.5% bisacrylamide and pour it into the beaker to a depth that will provide wells of sufficient volume to hold the protein sample. Allow the gel to polymerize. Remove the glass rods or test tubes.

   In this procedure, the intent is to prevent the protein from entering the gel. Alternatively, 2% to 5% agarose gels may be used.

   If the gels are to be stored, fill them with water to prevent drying. Remove the water and blot out any excess before adding a protein sample.

3. Add the protein sample to the well and let stand 30 min to 1 hr at room temperature.

   If the wells have a diameter >5 mm, dialyze for 1 to 2 hr or longer. The rate of dialysis is temperature dependent, so dialysis should be carried out at room temperature or higher if it does not adversely affect the protein.

4. Use a pipet to transfer the dialyzed protein sample from the well to a clean test tube suitable for the selected total protein assay.

TRICHLOROACETIC ACID PRECIPITATION OF PROTEIN SAMPLES

SUPPORT PROTOCOL 3

Trichloroacetic acid (TCA) precipitation can be used to precipitate proteins away from TCA-soluble, low-molecular-weight compounds that may interfere with assays for total protein. The procedure may also be used to concentrate protein from a dilute aqueous solution.

Add 10% (w/v) TCA to a protein sample to give a final concentration of 3% to 4% (v/v). Let stand 2 to 5 min at room temperature. Remove the supernatant and resuspend the precipitate in neutral buffer or alkali, depending on the method for further analysis.

If the sample is to be acid hydrolyzed (see Basic Protocol 4), carry out TCA precipitation in the heat-sealable tube, remove the supernatant, and dissolve the precipitate in 3% H_2SO_4.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

BCA reagent A

1 g 4,4′-dicarboxy-2,2′-biquinoline, disodium salt (Na_2BCA; Pierce or Sigma; 26 mM final)
2 g $Na_2CO_3 \cdot H_2O$ (0.16 M final)
160 mg disodium sodium tartrate (7 mM final)
0.4 g NaOH (0.1 M final)
0.95 g $NaHCO_3$ (0.11 M final)
100 ml H_2O

After mixing and dissolving the components, adjust the pH to 11.3 ± 0.2, using solid NaOH to increase pH or $NaHCO_3$ to decrease pH. Store this alkaline reagent in a plastic container 1 to 3 weeks at room temperature, longer at 4°C.

Only the disodium salt of Na_2BCA is soluble at neutral pH; the free acid is not readily soluble, even in alkali.

BCIT LIBRARY / CENTRAL UNIVERSITY LIBRARY

BCA reagent B

4 g $CuSO_4 \cdot 5H_2O$ (16 mM final)
100 ml H_2O
Store up to a few months at room temperature

BCA reagent A/reagent B mix

Mix 50 vol BCA reagent A with 1 vol BCA reagent B (see recipes). Prepare only sufficient mix for a few days' work because the mixture is not stable.

BCA reagent A/reagent B mix is light apple-green in color.

Biuret reagent

1.5 g $CuSO_4 \cdot 5H_2O$ (6 mM final)
6 g sodium potassium tartrate tetrahydrate ($C_4H_4KNaO_6 \cdot 4\ H_2O$, Rochelle salt; 21 mM final)
300 ml 10% (w/v) NaOH (0.75 M final)
Preboiled, cooled H_2O to 1 liter
Store at room temperature in plastic bottles

The water should be preboiled and cooled to remove dissolved CO_2.

Biuret reagent ordinarily is stable for several months. If a black color (due to cupric oxide) develops, discard the reagent.

CAUTION: *Discarded reagent and assay samples may be stored in open plastic buckets in a chemical fume hood to reduce volume (evaporate off water). The salts, especially copper, should be disposed of as hazardous waste.*

Coomassie dye reagent

100 mg Coomassie brilliant blue G250 [0.01% (w/v)]
50 ml 95% ethanol (5% final)
100 ml 85% phosphoric acid (8.5% final)
H_2O to 1 liter
Filter through Whatman no. 2 filter paper
Store up to 1 month at room temperature in a glass container

Coomassie brilliant blue G250 is available from Sigma (as brilliant blue G1), Bio-Rad, Pierce, and others.

Hartree-Lowry reagent A

2 g sodium potassium tartrate·4 H_2O (Rochelle salt; 7 mM final)
100 g Na_2CO_3 (0.81 M final)
500 ml 1 N NaOH (0.5 N final)
H_2 to 1 liter
Store 2 to 3 months at room temperature in a plastic container

Hartree-Lowry reagent B

2 g sodium potassium tartrate·4 H_2O (0.07 M final)
1 g $CuSO_4 \cdot 5H_2O$ (0.04 M final)
90 ml H_2O
10 ml 1 N NaOH
Store 2 to 3 months at room temperature in a plastic container.

Hartree-Lowry reagent C

Dilute 1 vol Folin-Ciocalteau reagent (Sigma) with 15 vol water.

Prepare this solution daily in 16-ml quantities or multiples thereof. Do not adjust the pH.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Ninhydrin reagent

Solution 1:

0.2 M citrate buffer, pH 5.0
21 g citric acid monohydrate (reagent grade; 0.22 M final)
200 ml 1 N NaOH (0.4 N final)
H_2O to 500 ml
Store at 4°C

Solution 2

0.8 g $SnCl_2 \cdot 2H_2O$ (7 mM final)
500 ml 0.2 M citrate buffer, pH 5.0 (Solution 1; 0.1 M final)

Solution 3

20 g ninhydrin, crystalline [Sigma; 4% (w/v) final]
500 ml 2-methoxyethanol (methyl cellosolve)

Working solution: Mix 500 ml Solution 2 and 500 ml Solution 3. Purge with nitrogen and store in dark bottles at 4°C.

The quantities of reagents can be scaled down.

Methyl cellosolve (2-methoxyethanol) often forms peroxides that need to be removed, e.g., by refluxing over tin metal and distilling. Alternatively, it can be purchased protected under nitrogen (e.g., from Aldrich). In any case, ninhydrin solutions in methyl cellosolve should be purged briefly with nitrogen after each use to exclude air.

CAUTION: *Methyl cellosolve has a boiling point of 124°C and is rather toxic, though not very volatile. Large quantities should be handled in a chemical fume hood.*

CAUTION: *Ninhydrin solution stains everything with an intense purple color. Always wear gloves and a lab coat when mixing and transferring ninhydrin solutions. Avoid spills. If skin becomes stained, wash with soap and water several times per day; color will persist for several days.*

COMMENTARY

Background Information

Many assays for quantitating total protein exist. Several are reliable and straightforward. How to chose the most suitable or optimal method is a recurring problem; the solution frequently requires the use of more than one method or protocol. A good strategy is to compare the results of two methods, such as A_{280} measurements and one of the copper-based chromogenic methods—assays that rely on different chemical properties. Very large differences in total protein estimates from two or more methods occur with crude preparations from bacterial cells, cell cultures, tissues, and food extracts, all of which may be laden with interfering substances. The principle question is often not how sensitive a particular protein assay happens to be, but rather how the assay is affected by interfering substances. The primary concern is how to maneuver around the interfering compounds or how to eliminate them altogether, using analytical protocols to follow the progress. Methods for dealing with some of the more common interfering substances are discussed below.

Individual total protein analytical methods commonly disagree with one another by as much as 5% to 20% even in the case of a well-behaved protein not laden with interfering compounds. Disagreement in the case of crude samples may be much greater. There are two general ways to deal with such discrepancies. The first is to remove interfering compounds that are likely to upset one, or perhaps both, analyses, or to resort to a third method, such as biuret analysis, if there is enough sample to resolve the differences. The second and ultimate is to conduct total nitrogen analysis, e.g., by the Kjeldahl method or modern versions of it. It is a generally reliable practice to assume that all proteins and polypeptides contain very close to 16.5% nitrogen by weight, so multiplication of the weight of nitrogen obtained from Kjeldahl analysis by a factor of 6.0 should provide a valid benchmark measure of the weight of protein in an ammonium salt–free sample.

The protocols and related commentaries in this unit describe specific, frequently used methods for quantitating total protein. How-

BOSTON CENTRAL UNIVERSITY LIBRARY

ever, this field is still evolving. *Analytical Biochemistry* frequently contains reports of new analytical methods, variants of existing methods, and methods for removing interfering compounds.

Recording and reporting data for total protein analysis

Calibration curves for spectrophotometric (colorimetric) assays for total protein are almost always plotted with the vertical axis in units of absorbance at the optimum wavelength (A_λ), which usually ranges between 0 and 1.00, or 0 and 2.0, depending on the spectrophotometer. The horizontal axis is usually labeled in units of protein such as µg or mg, often without specifying the related volume—the volume of sample introduced into the assay or the concentration per milliliter of sample or per milliliter of total assay volume. If a 0.50-ml sample is analyzed in a colorimetric assay that adds 3 to 6 ml of reagents, there is a >10-fold concentration difference involved. For this reason, it is very important to be clear and specific in defining the appropriate volume units for the values on the horizontal axis. Clear specification of units for calibration curves is also important because the numerical value of the slope is a direct measure of the sensitivity of the method and a reflection of the molecular composition of the calibrating protein(s). Molar absorption coefficients at specified wavelengths (ε_λ) are defined in Equation 3.4.3, where l is cuvette pathlength in cm, c is the molar concentration of the chromogen (moles liter^{-1}), and A_λ is the absorbance as read from the spectrophotometer. In most cases, path length $l = 1$ cm.

$$\varepsilon_\lambda = \frac{A_\lambda}{c \times l}$$

Equation 3.4.3

Because spectrophotometry-based total protein analyses produce calibration plots of A_λ versus concentration (c), the slopes of the plots are either directly equal to molar absorption coefficients (if c is in units of molarity) or directly proportional to molar absorption coefficients (if other units are used for the horizontal axis). A well-characterized pure protein (e.g., bovine serum albumin, BSA) used for calibration should produce a calibration curve with a slope (absorption coefficient) that is reasonable based on the composition of the protein for the chromogen or prochromogen. BSA has close to 18 phenolic (tyrosine) groups per protein molecule. Calibration of a procedure such as the Folin assay that depends on phenolic side chains should respond at least semiquantitatively to the protein's composition as well as its concentration variation. Folin assays for phenolic (and cresol) compounds produce ε_{700} values in the range 7500 to 8500 M^{-1} cm^{-1}; the Folin "color yield" for BSA should be within ±5% of the value for an equivalent number of cresol side chains.

Careful performance of total protein assays and recording of the results are helpful for two reasons. First of all, information from the calibration curve can be compared to known absorption coefficients for purified or partially purified proteins to evaluate the accuracy of the assay. Second, the results can be used as a guide in planning additional experiments to minimize the amount of sample protein required for analysis. Even approximate calculations of the amount of protein, numbers of tyrosine side chains in individual proteins, useful dilution ranges, or amounts of lyophilized protein to use can be helpful in the planning stages and initial benchwork.

It is not necessary for the horizontal axis to be plotted in units of molarity. However, it is desirable both for planning and for cross-checking results that units include a clear statement of volume so they can be readily converted to other units as needed for reporting and planning. One appropriate unit is µg or mg/ml of final assay volume (per ml f.a.v.), which specifies the concentration of the analyzed protein in the operational volume of a solution for spectrophotometric reading after the sample, reagents, and diluents have all been combined.

Table 3.4.2 lists the slopes of calibration curves, equivalent to absorption coefficients and assay sensitivities, for proteins and sugars in some commonly used assays. For some total protein assays, it is a good strategy to monitor sugars in parallel to evaluate removal of interfering compounds. Details for total sugar and reducing sugar assays are outlined in Lovrien et al. (1987) and references therein.

Proteins are frequently investigated and put into large-scale production and use before their molecular weights are known. However, their dry weights may be available. Hence the relationship between molar absorption coefficients (ε_λ) and weight absorption coefficients ($E^{1\%}$) at a stated wavelength is useful for converting data using Equation 3.4.2. The molar absorption coefficient and the weight absorption coefficient are related to absorbance readings at

RCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 3.4.2 Slopes of Calibration Plots for Spectrophotometric Assays[a]

Assay	Calibrating compound	Measured slope
Dinitrosalicylate (DNS)	Glucose	5.50 A_{575} (mg sugar/ml f.a.v.)$^{-1}$ cm^{-1}
Nelson-Somogyi reducing sugar	Glucose	6.3×10^{-3} A_{520} (nmol glucose/ml f.a.v.)$^{-1}$ cm^{-1}
Phenol–sulfuric acid neutral sugar	Mannose	8.6×10^{-2} A_{485} (µg sugar/ml f.a.v.)$^{-1}$ cm^{-1}
Biuret protein	BSA	2.3×10^{-4} A_{550} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}
Hartree-Lowry protein	BSA	1.7×10^{-2} A_{650} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}
Bicinchinonic acid protein	BSA	1.5×10^{-2} A_{562} (µg protein/ml f.a.v.)$^{-1}$ cm^{-1}
Colorimetric microkjeldahl nitrogen	Ammonium sulfate	1.3 A_{660} (µg nitrogen/ml f.a.v.)$^{-1}$ cm^{-1}

[a]Abbreviations: BSA, bovine serum albumin; f.a.v., final assay volume.

the stated wavelength λ (often, the wavelength of an absorbance maximum or peak), to concentrations, and to cuvette path length (l, in cm) as stated in Equations 3.4.3 and 3.4.4, where c

$$E^{1\%} = \frac{A_\lambda}{c \times l}$$

Equation 3.4.4

is concentration in dry g/100 ml.

Copper ion–dependent assays for total protein

The biuret, Lowry, and bicinchinonic acid (BCA) assays are all dependent on copper ions—cupric and cuprous ions. Prochromogenic reagents used in each copper-based assay—Folin reagent, Folin-Ciocalteau, and BCA reagent—depend on the extent of reduction of cupric ion (Cu^{2+}) to cuprous ion (Cu^{+}) to develop their color (Smith et al., 1985). The brilliant color that BCA reagent develops is the result of formation of a complex with Cu^{+} but not Cu^{2+}. In turn, the amount of Cu^{+} produced depends on the amount of protein present, so color development is a measure of the amount of protein (see Fig. 3.4.6).

Compounds that affect either cupric or cuprous ion chemistry will interfere with the assays. In the BCA assay, the reagent chelates the cuprous ion. In the biuret assay, alkaline cupric ion interacts with polypeptide chains. The biuret reaction which is based on the interaction of the cupric ion with the imide form of the polypeptide in strong base to ionize the proton off the polypeptide amide group, may be a precursor to more strongly chromogenic, later-developed color-producing reactions. Aliphatic amines and ammonia or ammonium ion are strong ligands for copper, at least for Cu^{2+}. Reducing agents such as glucose can certainly be expected to interfere with these assays. Proteins rich in disulfide and sulfhydryl groups such as keratin interfere because reduced sulfur compounds powerfully bind to and reduce Cu^{2+}. Cupric ions in general are avid coordinating metals for foreign compounds, precipitating with them or becoming reduced by them. Large concentrations of ammonium sulfate and some of the phosphates used in purification schemes interfere with cupric ion–based chemistries. A list of interfering compounds, many

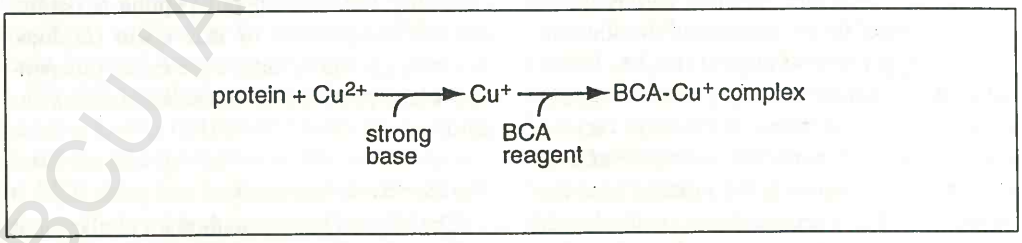


Figure 3.4.6 Reactions in the BCA assay.

BCU IASTM CENTRAL UNIVERSITY LIBRARY

Table 3.4.3 Interfering Compounds for Assays to Quantitate Total Protein[a]

Assay	Interfering compounds
Biuret	Ammonium sulfate Glucose Sulfhydryl compounds Sodium phosphate
Hartree-Lowry	EDTA Guanidine-HCl Triton X-100 SDS Brij 35 >0.1 M Tris Ammonium sulfate 1 M sodium acetate 1 M sodium phosphate
Bicinchinonic acid	EDTA >10 mM sucrose or glucose 1.0 M glycine >5% ammonium sulfate 2 M sodium acetate 1 M sodium phosphate
Acid digestion-ninhydrin	Ammonium sulfate Amino sugars
UV adsorption	Pigments Phenolic compounds Organic cofactors
Bradford	>0.5% Triton X-100 >0.1% SDS Sodium deoxycholate
Dry weigh	All buffer salts

[a]See Smith et al. (1985) for a more complete discussion.

of them used as buffers or found as metabolites in protein technology, is presented in Smith et al. (1985) and briefly summarized in Table 3.4.3.

Biuret assay

The biuret assay (see Basic Protocol 1) is periodically reinvestigated, using different formulations for the reagents and different proteins as calibrating standards. There are a number of modifications for the method scattered through the literature (Goshev and Nedkov, 1979). Some of these recommend the addition of KI, ~5 g per liter of reagent (Layne, 1957). Differences in apparent slopes from laboratory to laboratory may occur because of varying amounts of inert materials such as NaCl or water from incompletely dry proteins used for calibration. The average slope for the biuret assay, obtained using several globular proteins as calibration standards, is $\sim 2.5 \pm 0.3 \times 10^{-4}$ A_{550} (μg protein/ml of f.a.v.)$^{-1}$ cm^{-1}, about 100 times smaller than the Hartree-Lowry and BCA assays.

A perennial question in total protein analysis is the choice of calibration standard: should the protein be a purified sample of the protein under study or will BSA suffice? Although it is true that some total protein assays, notably UV-based A_{280} assays, are dependent on the amino acid composition of the protein, biuret chemistry is based on polypeptide structure, not the composition of side chain residues. Accordingly biuret assays are rather independent of the protein used as the calibration standard.

The Hartree-Lowry assay

The original Lowry method for total protein analysis (Lowry et al., 1951) has been reinves-

PCU IASC CENTRAL UNIVERSITY LIBRARY

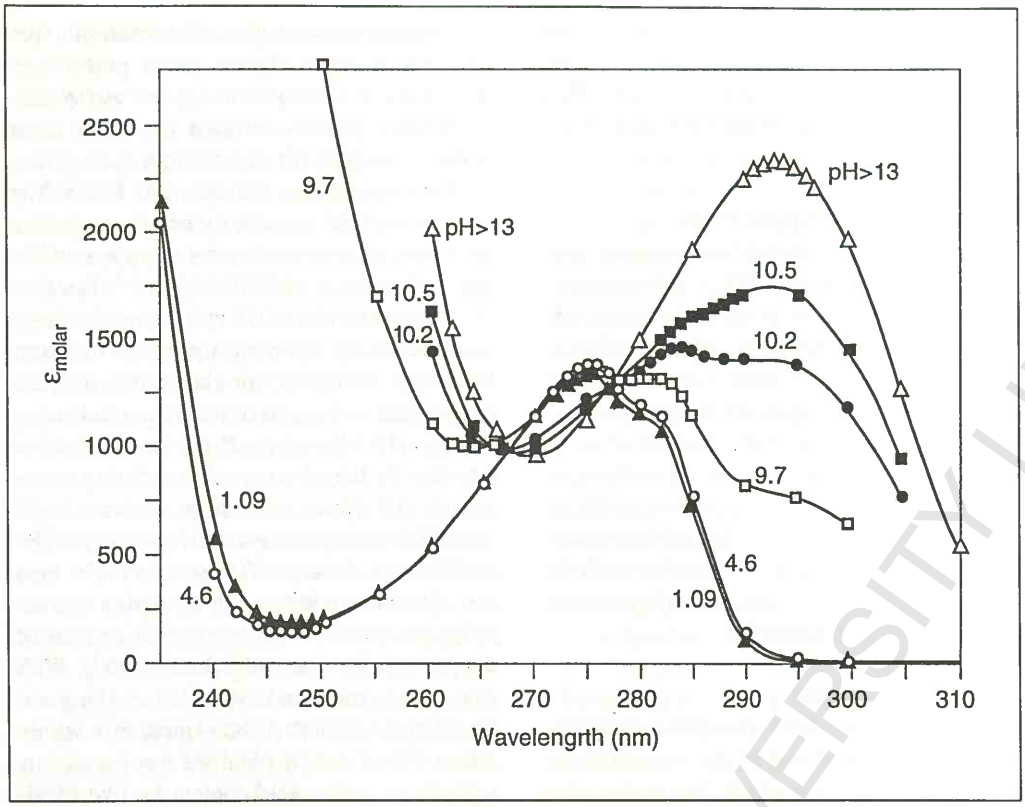


Figure 3.4.7 Absorption spectra of glycyl-L-tyrosine as a function of pH (Craig, 1967).

tigated many times to evaluate the effects of interfering compounds and the ability of detergents to solubilize otherwise insoluble proteins (reviewed by Peterson, 1983). The Hartree-Lowry assay (see Basic Protocol 2) yields linear results over a wider range of protein concentration, maintains the sensitivity, and is superior in the formulation of reagents (Hartree, 1972).

The bicinchinoinic acid assay

The bicinchinonic acid (BCA) assay (see Basic Protocol 3 and Fig. 3.4.6) developed by P.K. Smith et al. (1985) uses cupric ion in a biuret reaction with proteins in a strong base. The biuret reaction produces cuprous ion from cupric ion, and the cuprous ion is chelated with the BCA reagent. The Cu^{+}-BCA chelate is brilliantly colored with an absorption peak, λ_{max}, at 562 nm. Therefore, A_{562} is directly dependent on protein concentration, when the reaction conditions outlined by Smith et al. (1985; see Basic Protocol 3) are maintained.

The BCA method compares favorably with the older Lowry or Hartree-Lowry methods in sensitivity and convenience. The sensitivity of the BCA assay can be increased somewhat by using a higher concentration of the BCA reagent (i.e., of bicinchinonic acid disodium salt). However, the BCA reagent is rather expensive, and the fractional increase in sensitivity may not be worth the increase in cost.

The BCA assay is less susceptible than other assays to interference by a number of detergents. For this reason the BCA assay is sometimes favored with detergent-loaded samples, e.g., membrane and cellular proteins extracted by detergent solubilization. On the other hand, the BCA assay is susceptible to interference by reducing sugars, and even by ostensibly nonreducing sugars such as sucrose (because sucrose releases reducing sugars on partial hydrolysis; Spies, 1957). Some experimental systems use 10^{-3} to 10^{-4} M sulfhydryl reagents to protect proteins. These reagents can contain significant quantitites of reduced sulfhydryl reagents such as mercaptoethanol that react with cupric ion. Dialysis or other methods for removing organic –SH compounds from the sample may be required.

In order to avoid interfering substances in general, and also to increase protein concentrations, powerfully precipitating agents such as TCA (trichloroacetic acid) may be useful (Annand and Romeo, 1976; Beyer, 1983; also see Support Protocol 3). Most proteins in even dilute solution are quantitatively precipitated by TCA (at a concentration of a few percent),

DUKE UNIVERSITY CENTRAL UNIVERSITY LIBRARY

which also concentrates them. The supernatant is discarded, which rids the sample of most low-molecular-weight impurities. The TCA precipitate should be redissolved in base for subsequent BCA assay.

Acid digestion–ninhydrin assay

The acid digestion/ninhydrin assay (see Basic Protocol 4; Rosen, 1957) takes considerably more time than other methods that depend on starting with protein samples in solution. This assay is useful for dried solid samples such as plant material because the whole sample is hydrolyzed under conditions that convert peptides and proteins to amino acids suitable for conventional chromatographic analysis (Marks et al., 1985). Nucleic acid bases such as adenine apparently do not interfere with the ninhydrin reaction, which can also be advantageous with complex samples.

UV spectrum analysis

The ultraviolet (UV) absorption spectrum for a sample protein can be used to quantitate the protein and to evaluate the purity of a sample. Figure 3.4.7 shows the UV absorption spectra for glycyl-L-tyrosine. The molar absorption coefficient ε as a function of wavelength λ is dependent on pH. The glycyl moiety has negligible effect on the spectra, so they are close in general shape to those of tyrosine side chains in proteins. Tryptophan side chains (indole groups) have similar spectra at acid pH. However, tryptophan absorption is considerably more intense at 275 to 280 nm, by a factor of nearly 4. Thus, although different proteins have very variable tyrosine-tryptophan compositions, the general shape of their UV absorption spectra in acid are fairly close to those of tyrosine at neutral and acid pH.

Proteins in neutral and acid solutions exhibit a deep trough in absorption at 245 to 250 nm Most proteins, except for heme proteins and proteins with cofactors such as NADH, do not absorb from either tyrosine or tryptophan above 300 to 310 nm. Increased absorption at ≤250 nm and/or ≥300 nm is a strong indication of the presence of compounds that interfere with A_{280} measurements and their interpretation. The effectiveness of techniques such as dialysis and size-exclusion chromatography for removing interfering UV-absorbing compounds can be monitored by analyzing the sample for absorption at 250 nm (trough), 280 nm (peak), and 300 to 310 nm (residual). Ratios of these absorbances are often used to determine whether the UV spectrum is that of pure or contaminated protein. For example, the peak/trough ratio (A_{280}/A_{250}) for proteins in neutral and acid solutions is almost always ≥2.

When a protein solution is shifted from acidic to alkaline pH, the absorption spectrum for tyrosine changes radically due to alkaline ionization of the tyrosine hydroxyl group. The maximum change in absorption occurs at 295 nm, with a $\Delta\varepsilon_{295}$ of 2470 M^{-1} cm^{-1}. Proteins absorb rather little at 295 nm, except for some contribution by tryptophan groups. However, titration of trytophan side chains does not produce a shift in A_{295}, so observation of changes in A_{295} with changing pH can be ascribed to tyrosine in the absence of interfering compounds that behave similarly to tyrosine. Analysis of UV absorption spectra, especially at 280 and 295 nm, for a purified protein under acid and alkaline conditions can be used to characterize the tyrosine and tryptophan content of the protein (Bencze and Schmid, 1957). With care, both tyrosine and tryptophan can be quantitated by the Bencze-Schmid method to within 3% to 4% of values obtained from complete analysis of amino acid content by hydrolysis and chromatography (Stein and Moore, 1954; Chang, 1992).

A number of sophisticated UV spectrum–based methods for protein quantitation and characterization use absorption at 205 nm and 224 to 236 nm (Bencze and Schmid, 1957), and second-derivative spectra of enzymatic digests of proteins (Bewley, 1982). See *UNIT 3.1* for additional information and discussion. It is necessary to remove oxygen from the spectrophotometer by nitrogen purging for measurements at very low wavelengths ("hard" UV, <220 nm).

All proteins absorb strongly at 205 nm, so at first glance this seems an attractive means for quantitating protein. However, most published methods describing hard UV absorbance are based on the behavior of pure, mostly single-subunit proteins. Buffers with carbonyl groups and many other impurities also become strongly absorbing in the hard UV range, and the problems of conventional UV absorbance are magnified in the hard UV range.

Coomassie dye binding assay (Bradford assay)

Coomassie blue dye (see Basic Protocol 6) produces a rather large absorption spectrum shift in the visible-light range when it binds to many globular proteins. The strong binding forces are partly, but not entirely, due to electrostatic attraction between dye molecule sulfonate groups (two on the G250 dye molecule)

BCU IASRI LIBRARY UNIVERSITY COPYRIGHT

and cationic groups of proteins when they are made quite acidic. The pH of the reaction mixture is low enough to thoroughly convert proteins to cations, but not so low that any sulfonate groups can become neutralized. The low pH of ~0.8 M phosphoric acid also causes most proteins to expand so that they become exposed to the triphenylmethane, organic, and nonpolar sectors of the dye to enhance binding (Craig, 1967). Proteins that do not bind anionic dye, even in very low pH, may not bind Coomassie dyes. Early research on Coomassie dye–based protein assays used globular proteins, which are known to avidly bind dye and other similar anions such as detergent anions (Edsall and Wyman, 1958). Albumin, hemoglobin, chymotrypsinogen, and cytochrome *c*, which Bradford used to calibrate the method (Bradford, 1976), are particularly able proteins in this regard. However, other classes of proteins are quite unlike these globular, organic anion–binding proteins and these differences may be at the root of some of the failures reported for the Bradford method (Pierce and Suelter, 1977). At any rate, the Bradford method is especially simple and has good sensitivity when it is calibrated with proteins to which it is sensitive (e.g., albumin).

Jernejc et al. (1986) compared four different methods for determining total protein—Kjeldahl nitrogen, biuret, Lowry, and Coomassie dye binding—for each of eight stages in the isolation of proteins from mycelia. They found that the Coomassie dye binding method often grossly underestimated protein content, by factors >2-fold, relative to the other methods tested. Although Coomassie dye binding is an easy and convenient assay, its reliability must be verified for each experimental system.

A similar method is the Udy dye binding method, which uses acid orange 12 dye, a sulfonated azobenzene dye. The techniques was originally developed for food proteins, and it apparently compares well with Kjeldahl nitrogen analytic criteria (Udy, 1956).

Critical Parameters

There are two endemic problems in total protein assays. First, there is the issue of interfering compounds and how to remove them or minimize their effects. Second is the intense barrage of advertisements for reagents for measurement of total protein; these promise simplicity, sensitivity, and by implication suggest there is little need to consider what interfering compounds may do, how they operate, or in what concentrations they operate. In crude proteins and at many stages in protein purification, interfering compounds often are present in concentrations one, two, or even three orders of magnitude larger than protein concentrations. Some of the well-advertised means for measuring total proteins are based on calibration with pure proteins such as crystalline bovine serum albumin, which can be expected to behave well. However, good performance in calibration with single, fairly pure proteins frequently does not translate into reliable performance with real samples—tangled mixtures of carbohydrates, lipids, and nucleic acids with proteins and glycoproteins. In some cases, when samples are crude, approximately half the effort needs to be devoted to controls and to simple steps that can be taken to either eliminate interference or at least help understand it. Support Protocols 2 and 3 describe two techniques—dialysis and TCA precipitation—for removing low-molecular-weight interfering compounds.

It should be noted that calibrating a total protein assay using a pure protein, usually bovine serum albumin, as a reference standard is common practice. However, that does not ensure that data from crude samples, taken through the same procedures, accurately quantitates total protein in crude samples.

Variability in BCA analysis

There are a number of factors that contribute to variability in the BCA assay: the temperature and duration of the incubation step just before spectrophotometry, the use of too much protein, the choice of calibration protein, and the presence of interfering compounds in samples of unknown character. It may be helpful to experiment with the temperature and duration of incubation. Sometimes incubating the sample 30 min at 60°C can accelerate color development without loss of color stability (Smith et al., 1985; Goldschmidt and Kimelberg, 1989). However, incubated samples must be cooled adequately before making spectrophotometric measurements, because residual temperature gradients in the cuvettes can cause refractive index gradations or striations, which distort the optical path producing anomalies and errors. These errors can be neutralized by remixing samples until their temperatures are close to the temperature of the cuvette and they appear uniform by simple visual inspection.

BCA calibration plots (see Fig. 3.4.3) show definite curvature downward for most proteins in excess of ~90 to 100 µg protein/ml final assay volume. Such nonlinearity can occur

UCSD ASHOKA TRAN UNIVERSITY LIBRARY

when the protein or polypeptide are present in excess with respect to the amount of reagents (Cu^{2+} and BCA) available to develop color.

There is considerable variation in the slope of calibration plots and therefore in apparent sensitivities, depending on the protein used as a calibration standard. For conventional proteins—e.g., BSA, chymotrypsin, and ribonuclease—slopes can vary by as much as a factor of two (Smith et al., 1985). If gelatin is used for calibration with an incubation of 30 min at 37°C, the variation can be up to 3-fold. Some preparations of proteins, even "crystalline" and lyophilized proteins, still contain appreciable amounts of water and/or salts used in isolation or crystallization. Water contents may be measured by the dry weight method (see Basic Protocol 7). Some freeze-dried preparations are actually only 95% protein or even less. Accordingly, water and other contaminants in supposedly dried proteins may contribute to lowered slopes, sensitivities, and apparent absorption coefficients.

UV absorption

UV absorption spectra are useful for quantitating and characterizing pure proteins (also see *UNIT 3.1*). However, UV absorption is probably the most profoundly and easily distorted analytic method of quantitating proteins because concentrations of UV-absorbing contaminants as low as 10^{-4} to 10^{-5} M can affect the spectra. Thus caution is necessary when analyzing protein samples that are not very pure. Pigments, oxidation products, tannin-like substances, Maillard and other condensation products—all manner of biochemical compounds and natural products—all absorb at ultraviolet wavelengths, sometimes strongly, potentially causing very large errors (at least an order of magnitude). If a protein sample absorbs UV above ~320 nm, it is likely that appreciable and perhaps unacceptably large amounts of interfering compounds are present. Elaborate means have been proposed by a number of authors for UV analysis of total protein (reviewed by Peterson, 1983); nearly all are built on the supposition that no UV-absorbing interfering compounds are present in the protein sample. In practice, this supposition is very shaky, particularly at early and intermediate stages in protein purification of proteins. It may be valid for rather well-purified proteins. In general, UV absorption should be used in conjunction with other methods for analysis of total protein to correlate results. Additionally, UV absorption spectra can be used in a reverse sense, in conjunction with other analyses, to estimate the amounts of absorbing impurities (Beyer, 1983).

TCA precipitation

Trichloroacetic acid precipitation may also be affected by reagents used to prepare the sample. Chang (1992) describes a procedure for TCA precipitation of a solution containing large amounts of the strong anionic detergent sodium dodecyl sulfate (SDS). Sometimes it is useful to precipitate the interfering compounds leaving the proteins in solution. Polyethyleneimines (PEI; Sigma) can be used to precipitate nucleic acids, which have a UV absorption peak at 260 nm (Jendrisak, 1987). PEI does not absorb in the ultraviolet range. However, neither TCA nor PEI is free of cross-precipitation, so it may be necessary to experiment with the concentration of precipitant, pH, and temperature to optimize a particular precipitation step.

Anticipated Results

Table 3.4.4 presents a summary of the detection ranges to expect for the assays described here.

Table 3.4.4 Summary of Assays for Quantitating Total Protein

Assay	Sample size	Detection range (μg/ml)
Biuret	1 ml	1000–10,000
Hartree-Lowry	1 ml	100–600
BCA	100 μl	200–1000
Ninhydrin	1 ml	20–50
UV absorption	1 ml	30–300
Coomassie dye binding	100 to 200 μl	60–300
Dry weight	0.5 to 10 ml	2000–10,000

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Time Considerations

Basic Protocols 1, 2, and 3 (biuret, Hartree-Lowry, and BCA assays) require one hour for measuring sample, mixing, incubation, and reading the absorbance. Basic Protocol 4 (acid digestion–ninhydrin detection) requires one hour on the first day to measure the sample and seal the tubes, followed by an overnight hydrolysis. On the second day, 1 to 2 hr are required for opening the tubes, neutralization, mixing reagents, heating, and reading the resulting color. Basic Protocol 5 (UV absorption) is the simplest of the methods described in this unit, taking only about 30 min to make dilutions, warm up the spectrophotometer, and measure the absorbance curve. Basic Protocol 6 (Coomassie dye binding) is also fast, requiring 20 to 30 min. Basic Protocol 7 (dry weight determination) requires a long drying time (4 hr to overnight) plus 10 to 20 min each time the container is weighed.

Literature Cited

Annand, R. and Romeo, P.L. 1976. Protein in foodstuffs. *Amer. Lab.* 8(10):37-46.

Bencze, W.L. and Schmid, K. 1957. Determination of tyrosine and tryptophan in proteins. *Anal. Chem.* 29:1193-1196.

Bewley, T.A. 1982. A novel procedure for determining protein concentrations from absorption spectra of enzyme digests. *Anal. Biochem.* 123:55-62.

Beyer, R.E. 1983. A rapid biuret assay for protein of whole fatty tissues. *Anal. Biochem.* 129:483-485.

Bradford, M.M. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.

Chang, Y.-C. 1992. Efficient precipitation and accurate quantitation of detergent-solubilized membrane proteins. *Anal. Biochem.* 205:22-26.

Craig, L.C. 1967. Techniques for the study of peptides and proteins by dialysis and diffusion. *Methods Enzymol.* 11:870-884.

Edsall, J.T. and Wyman, J. 1958. Biophysical Chemistry. pp. 591-594. Academic Press, New York.

Freifelder, D. and Better, M. 1982. Dialysis of small samples in agarose gels. *Anal. Biochem.* 123:83-85.

Goldschmidt, R.C. and Kimelberg, H.K. 1989. Protein analysis of mammalian cells in monolayer culture using the bicinchinonic acid assay. *Anal. Biochem.* 177:41-45.

Goshev, I. and Nedkov, P. 1979. Extending the range of application of the biuret reaction: Quantitative determination of insoluble proteins. *Anal. Biochem.* 95:340-343.

Hartree, E.F. 1972. Determination of protein: A modification of the Lowry method that gives a linear photometric response. *Anal. Biochem.* 48:422-427.

Jendrisak, J. 1987. Use of polyethyleneimine in protein purification. *In* Protein Purification: Micro to Macro (R. Burgess, ed.) pp. 75-97, Alan R. Liss, New York.

Jernejc, K., Cimerman, A., and Perdih, A. 1986. Comparison of different methods for protein determination in *A. niger* mycelium. *Appl. Microbiol. Biotechnol.* 23:445-448.

Kirazov, L.P., Venkov, L.G., and Kirazov, E.P. 1993. Comparison of the Lowry and Bradford protein assays applied for protein estimation of membrane-containing fractions. *Anal. Biochem.* 208:44-48.

Layne, E. 1957. Spectrophotometric and turbidimetric methods for measuring proteins. *Methods Enzymol.* 3:447-454.

Lovrien, R., Goldensoph, C., Anderson, P.C., and Odegaard, B. 1987. Three Phase Partitioning (TPP) via *t*-butanol: Enzymes separation from crudes. *In* Protein Purification: Micro to Macro (R. Burgess, ed.) pp. 131-148. Alan R. Liss, New York.

Lovrien, R.E., Conroy, M.J., and Richardson, T.I. 1995. Molecular basis for protein separations. *In* Protein-Solvent Interactions (R.B. Gregory, ed.) pp. 521-553. Marcel Dekker, New York.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193:265-275.

Marks, D.L., Buschbaum, R., and Swain, T. 1985. Measurement of total protein in plant samples in presence of tannins. *Anal. Biochem.* 147:136-143.

Peterson, G.L. 1983. Determination of total protein. *Methods Enzymol.* 91:95-119.

Pierce, J. and Suelter, C.H. 1977. An evaluation of the Coomassie brilliant blue G-250 dye-binding method for quantitative protein determination. *Anal. Biochem.* 81:478-480.

Rosen, H. 1957. A modified ninhydrin colorimetric analysis for amino acids. *Arch. Biochem. Biophys.* 67:10-15.

Sedmak, J.J. and Grossberg, S.E. 1977. A rapid, sensitive, versatile assay for protein using Coomassie Brilliant blue G250. *Anal. Biochem.* 79:544-552.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. 1985. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150:76-85.

Sober, H.A. (ed.) 1970. Molar extinction coefficients and $E^{1\%}$ values for proteins at selected wavelengths of the ultraviolet and visible region. *In* Handbook of Biochemistry, Selected Data for Molecular Biology pp. C71-C98. Chemical Rubber Co. Press, Cleveland.

BOSTON UNIVERSITY LIBRARY

Spies, J.R. 1957. Colorimetric procedures for amino acids. *Methods Enzymol.* 3:467-471.

Stein, W.H. and Moore, S. 1954. The free amino acids of human blood plasma. *J. Biol. Chem.* 211:915-926.

Udy, D.C. 1956. A rapid method for estimating protein in milk. *Nature* 178:314-315.

Van Kley, H. and Hale, S.M. 1977. Assay for protein by dye binding. *Anal. Biochem.* 81:485-487.

Watters, C. 1978. A one-step assay for protein in the presence of detergent. *Anal. Biochem.* 88:695-698.

Contributed by Rex Lovrien and Daumantas Matulis
University of Minnesota
St. Paul, Minnesota

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Kinetic Assay Methods

The design and use of a particular enzyme assay is dependent on the objectives of the particular investigation. If the focus is on enzyme purification or issues of clinical biochemistry, experimental protocols typically require optimizing conditions for all substrates and effectors—e.g., with respect to pH value, temperature, buffering agent, and ionic strength. However, optimal conditions may not be reasonable for all parameters and factors. For example, substrate instability may preclude studies at the enzyme's optimal pH or at the physiological temperature. Thus, Silverstein and Sulebele (1969) examined the exchange rates of malate dehydrogenase at 1°C to avoid problems arising from the decomposition of oxaloacetate. Likewise, in clinical studies and enzyme purification, saturating concentrations of substrate(s) are typically required, although this may not be possible or even desirable in many instances, as limitations in substrate solubility or substrate inhibition of the enzyme may discourage the use of elevated concentrations. The focus in clinical studies and enzyme purification is on providing good estimates of the activities of a specific enzyme in a number of physiological samples with a rapid, straightforward, and reproducible assay method. However, for the kineticist the objective is to characterize the initial rates of the enzyme-catalyzed reaction under a wide set of conditions. It is through these detailed studies, often carried out at subsaturating substrate concentrations, that the investigator is able to determine the true optimal parameters to be used in protein purification and clinical studies, as well as to probe the dynamics and mechanism of the reaction and its modulation by other factors. The purpose of this unit is to provide a brief review of issues important in the design of initial-rate assay methods.

In most instances, the investigator already has some information about the kinetic properties of the enzyme, obtained from preliminary experiments as well as from the literature. It is necessary to have an approximation of the magnitude of the Michaelis constants of all the substrates in order that detailed initial rate studies can be effectively designed. Certainly, information on the identity of the substrates and products is required, along with an effective procedure for monitoring substrate depletion or product accumulation. The nature of the reaction and the properties of the substrates help dictate the assay procedure used. Enzyme-assay methods include those based on spectrophotometry, radiometry, pH-stat, turbidometry, polarography, and fluorometry. Spectrophotometric and fluorometric assay protocols rely on the differences in spectral or fluorescent properties, respectively, of the product(s) relative to the substrate(s). With radiometric procedures, effective means for separating labeled substrate(s) from labeled product(s) must be available so that the amount of product formed at a given time can be accurately measured. Those reactions that utilize or generate protons (or hydrons) can be followed by pH-stat methods. If the reaction studied results in a change in solution turbidity (e.g., caused by aggregation of a macromolecular complex), then turbidometric protocols can effectively assay for the protein that causes the change in turbidity. All of these methods, as well as others, can be found in Bergmeyer (1983).

GENERAL ASPECTS OF ASSAY DESIGN

An enzyme's progress curve (i.e., a plot of product formation or substrate depletion as a function of time; see Fig. 3.5.1) consists of four phases that are of interest to the investigator. The duration of the **pre-steady-state** region is relatively short and is typically followed via rapid-flow or relaxation techniques. Discussion of this region of the progress curve is beyond the scope of this unit, but a review of these techniques has recently appeared (Fierke and Hammes, 1995). The next portion of the curve is the **steady-state** phase, a longer phase characterized by relatively constant concentrations of each of the individual enzyme forms (e.g., the enzyme-substrate binary complexes). During the third phase of the progress curve—the **post-steady-state** region—the levels of these enzyme complexes change rapidly, until the last phase—**equilibrium**—has been reached. Each of these phases contains information of use to the investigator. The steady-state portion of the curve typically exists for a few to several hundred seconds. It is this region of the curve that is most accessible to the majority of investigators. In addition, the relatively straightforward mathematical treatment of steady-state kinetics provides an excellent foundation for the subsequent characterization of the enzyme system. Because enzyme rate expressions can be quite complex, a number of

Copyright © 1996 by John Wiley & Sons, Inc.

BIJVANK CENTRAAL UNIVERSITY LIBRARY

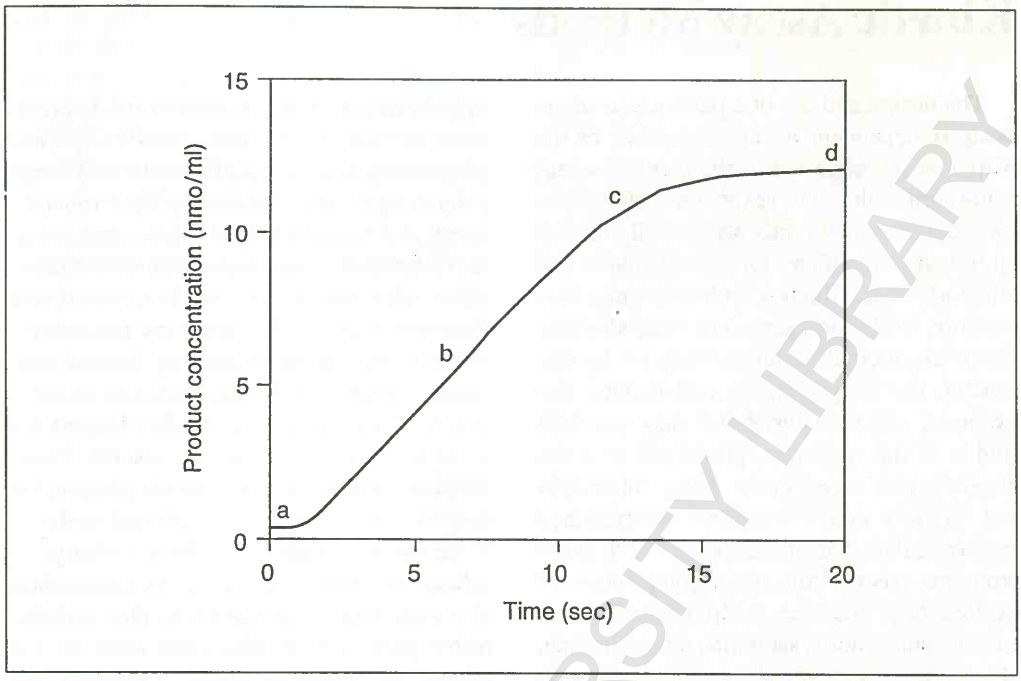


Figure 3.5.1 Typical progress curve for an enzyme-catalyzed reaction, where *a* represents the pre-steady-state region, *b* the steady-state region, *c* the post-steady-state region, and *d* the equilibrium region.

assumptions are made to linearize the equations. The investigator should ensure that each of these factors are addressed when any assay protocol is being designed.

Maintaining Constant Reaction Conditions

The status of each experimental parameter—e.g., temperature, ionic strength, and pH—should remain constant over the time course of the assay.

Temperature

Prior to initiation of the reaction by addition of a small aliquot of enzyme or substrate, the reaction mixture should be thermally equilibrated with an accurately controlled constant-temperature water bath for several minutes. If the substrate is thermally stable, this incubation should be for ≥10 min or until the desired reaction temperature is reached. If the substrate is unstable at this temperature, it may be necessary either to assay at a different temperature, to correct for substrate decay with time, or to initiate the reaction by addition of a small aliquot of substrate solution that was freshly prepared and maintained in an ice bath. When the reaction is initiated by addition of either the enzyme or the substrate, the volume of the aliquot should be small enough to minimize any alterations in temperature upon addition. Typically, there is at least a hundred-fold difference in volume between the reaction mixture and the initiating aliquot. If the enzyme and substrate are stable under the assay conditions, then both solutions should be preincubated at this temperature.

It should be emphasized that valid initial rate experiments are only obtained when good temperature control is maintained. The temperature-regulated water bath should have a variability of ≤0.1°C. When the reaction is initiated, the solution should be thoroughly mixed. If reaction progress is followed using an instrument such as a spectrophotometer, the sample compartment should be thermally isolated.

Another issue of concern is choice of the temperature at which enzyme activity will be measured. In 1964, the International Union of Biochemistry (IUB) recommended a temperature of 30°C. Clinical chemists have often suggested a change to 37°C. However, the final choice of temperature will be dependent on the system being investigated. For example, the activity of glutamate dehydrogenase isolated from the hyperthermophile *Pyrococcus furiosus* is negligible at both 30°C and 37°C; the enzyme exhibits maximal activity at 97°C (Klump et al., 1992). Clearly, at some point in the investigation, the effect of temperature on the catalytic activity will have to be characterized.

PCU IASL CENTRAL UNIVERSITY LIBRARY

Buffer selection

In setting up the assay protocol, care should also be exercised in selecting an appropriate buffer agent. Unfortunately, many investigators choose a buffer solely on the basis of the desired pH of the reaction medium. Many excellent buffers, however, have a significant oxidation-reduction potential that can severely alter an enzyme's activity. For example, cacodylate ($pK_a = 6.27$) and other organoarsenicals have a potent oxidizing potential under acidic conditions. Borate ($pK_a = 9.23$) and other borate-based buffers form complexes with many diols, polyols, carbohydrates, and ribonucleotides; for example, in the presence of borate ion, both L- and D-serine form a transition complex and strongly inhibit γ-glutamyl transpeptidase (Tate and Meister, 1978). Hence, care should be exercised in selecting a buffer.

In determining which buffer is optimal for a given enzyme system, it is best to examine the enzyme's activity with as many buffers as possible. Buffers in whose presence the enzyme exhibits high activity should then be examined to determine if the buffering agent interacts with the protein. If pH and ionic strength are held constant, the enzyme activity should be independent of the buffer concentration. Another factor of importance in selecting the appropriate buffer is the capacity of the buffering agent to bind metal ions. Ideally, a good buffer will not form any significant complexes with metal-ion cofactors.

Metal-ion concentrations

The role of metal ions is crucial with many enzymes, particularly with systems that require nucleotides—e.g., phosphotransferases. Care must be exercised in designing initial-rate assays to ensure that the concentration of free metal ions does not vary significantly with substrate concentration. Thus, when varying the concentrations of substrates such as ATP, the concentration of free, uncomplexed Mg^{2+} should remain relatively constant. One of two methods is commonly utilized to achieve this end: either (1) maintaining the total nucleotide concentration and the metal concentration at a constant stoichiometric ratio (e.g., 10:1) or (2) maintaining a constant excess of metal ion over the total nucleotide concentration (e.g., 1.0 mM free Mg^{2+}). The second method, which has been demonstrated to be the preferable of the two (O'Sullivan and Smithers, 1979), requires knowledge of the stability constant for the metal-ion–nucleotide complex under the conditions of the assay.

Substrate Stability and Purity

It is crucial to establish the degree of stability of the substrate as well as that of any enzyme effector under the assay conditions (also see discussion of Maintaining Constant Reaction Conditions). Substrate and effector stabilities can be tested by preparing a solution of the compound at a concentration that will be used experimentally with the other components of the assay mixture, but without the enzyme. Then, samples are removed at different periods of time, usually over a period of several hours, and assayed either chemically or enzymatically (Bergmeyer, 1983). A plot of concentration as a function of time will reveal if the substrate is stable or if there is a decay. When present, this decomposition is usually a first-order process and can be described by the equation $[S]_t = [S]_0 e^{-kt}$, where $[S]_t$ and $[S]_0$ are the substrate concentrations at time t and time 0, respectively, and k is the first-order rate constant. If the decomposition is greater than first-order or composed of more than one first-order process, then the decay must be evaluated by the appropriate expression, which can be found in any basic kinetics text.

An issue related to substrate stability is substrate purity. Impurities in the substrate, effectors, or enzyme preparation can be potential sources of experimental error. Because alternative substrates and competitive inhibitors are often structurally similar to the substrate and frequently have similar chemical and physical properties, it is not surprising to find these compounds as impurities in commercial substrate preparations. Spectral and chromatographic analyses, as well as enzymatic analysis, are crucial in establishing the purity of a particular biochemical substance. If impurities are present in a reaction mixture, a kinetic analysis will give wrong estimates of the kinetic parameters (Allison and Purich, 1979). Heavy-metal ions are often found as impurities in commercial substrates. To remove these, it may be necessary to recrystallize the substrates, possibly in the presence of a chelating agent such as EDTA.

These issues of purity and stability are also pertinent for any modulators of enzyme activity, including the product(s). The products have to be stable over the time course of the assay. In some instances, the precise nature of the true product of the enzyme-catalyzed reaction may not be known. In these cases, the investigator may follow the reaction via substrate depletion. It is frequently useful to do a kinetic analysis of freshly prepared substrate in the absence and

FOR INSTRUCTIONAL USE ONLY

presence of enzymatically prepared product. A decent understanding of the possible presence of product inhibition in initial-rate experiments is always useful to the protein chemist.

Enzyme Stability and Purity

Enzymes often undergo a slow inactivation. Usually, the rate of this inactivation is accelerated by dilution of the stock enzyme preparation. Efforts have to be made to eliminate or lessen the loss of enzyme activity with time. Often, multisubstrate enzymes are stabilized by the presence of a substrate. For example, tubulin:tyrosine ligase is stabilized by the presence of its protein substrate (Deans et al., 1992) and the presence of Mg^{2+}–ATP is known to stabilize a number of phosphotransferases. If the enzyme inactivation can be retarded by the presence of a substrate, the investigator should note the amount of that substrate that is present with the enzyme so that proper corrections can be made when the reaction is initiated with the full assay mixture.

Other agents—e.g., glycerol, certain salts, high-ionic-strength solutions, and EDTA—have also been used to stabilize proteins. Often, inactivation is a result of surface denaturation or adsorption of the enzyme on glass surfaces following dilution. The presence of stabilizing agents and/or use of plastic containers can lessen these effects significantly. The presence of serum albumin or another protein will frequently minimize the effects of enzyme dilution; the investigator should test several proteins when this technique is used. The stabilizing protein may also bind the substrates or the enzyme's cofactors; if this is the case, corrections have to be made.

If a stabilizing agent cannot be identified, the inactivation should be minimized by using the smallest dilution possible. If practical, the reaction may be initiated by the addition of one of the substrates. Corrections for loss of enzyme activity can also be made by establishing an activity-decay curve. However, it should be verified that the loss of activity is due to total inactivation and not simply to a conformational change to a form having less activity (i.e., a form with different kinetic parameters). This possibility could be tested by kinetic assays at significantly different times following dilution or by two reference assays, one using a saturating substrate level and the other using a subsaturating substrate level.

The presence of thiol reagents—e.g., 2-mercaptoethanol, dithiothreitol, or dithioerythritol—is routinely helpful in maintaining activity for those enzymes that require reduced thiol groups for activity. Care must be exercised that the reducing agent does not produce other effects. For example, γ-glutamylcysteine synthetase is inhibited by the presence of thiol reagents even though cysteine is a substrate (Seelig and Meister, 1985).

Enzyme purity is also an issue of concern. It is important in early initial-rate studies to eliminate contaminating activities and to demonstrate that any minor activity does not substantially affect the reaction being studied. For multisubstrate enzymes, this can be readily accomplished by observing the stability of each substrate, effector, and product separately in the presence of the enzyme preparation. Adenylate kinase is a common contaminant in preparations of a wide variety of phosphotransferases and other nucleotide-dependent enzymes. This contaminating enzyme can be inhibited by performing the assay at high ionic strength (Bowen and Kerwin, 1955). Addition of elevated levels of AMP will also inhibit adenylate kinase. However, many phosphotransferases are also inhibited by this mononucleotide or are themselves inhibited by high ionic strength. Purich and Fromm (1972) and Lienhard and Secemski (1973) have shown that the multisubstrate analogs P^1,P^4-di(adenosine-5′)-tetraphosphate and P^1,P^5-di(adenosine-5′)-pentaphosphate are potent and specific inhibitors of this activity.

In some instances it may be impossible to completely eliminate side reactions. Enzymes that have branched reaction pathways (i.e., multifunctional enzymes) will exhibit two pathways. Most multifunctional enzymes are transferases having multiple acceptor specificity. Glucose-6-phosphatase catalyzes not only the hydrolysis of glucose-6-phosphate but also its synthesis; thus, it may effect the transfer of phosphoryl groups from pyrophosphate, carbamoyl phosphate, or other glucose-6-phosphate molecules (Nordlie, 1982). This synthetic role of glucose-6-phosphatase is not a minor activity; at a concentration of 100 mM D-glucose, the synthesis and hydrolysis activities are approximately equal. The enzyme γ-glutamyl transpeptidase catalyzes the hydrolysis of a number of γ-glutamyl-containing molecules (including glutathione and glutathione disulfide) and also exhibits a transferase activity in which the glutamyl moeity is transferred to any of a number of amino acid acceptors, particularly cysteine, glutamine, and methionine (Allison, 1985). The hydrolase and transferase activities of γ-glutamyl transpeptidase

BCU IASE CENTRAL UNIVERSITY LIBRARY

are approximately equal in the presence of a pool of acceptor substrates at their physiological concentrations (Allison and Meister, 1981). Asparagine synthetase catalyzes the glutamine-dependent synthesis of L-asparagine and also exhibits a significant glutaminase activity (Boehlein et al., 1994). In all of these cases, it is possible to use inactivation studies to observe the parallel loss of both (or all) activities with time.

Enzyme Concentration

The duration of the steady-state phase of an enzyme-catalyzed reaction is dependent on a number of factors. Perhaps the most important of these is the relative concentration of the enzyme and substrate. The ratio of the initial substrate concentration to the total enzyme concentration should be ≥100:1. With lower values, the steady-state phase will be of very short duration. It should be pointed out that this relationship holds for enzyme effectors as well. Thus, when investigating the influences of an activator or competitive inhibitor on enzyme activity, the total enzyme concentration should be much less than the concentration of the effector. This may prove difficult if the effector has a very tight affinity with the enzyme. In these cases, the protocols outlined by Williams and Morrison (1979) will prove useful.

Initial Rate and Steady-State Condition

In an experiment, the investigator typically measures the rate of an enzyme-catalyzed reaction by calculating the slope of the progress curve. This procedure is only valid if the plot of product concentration as a function of time has been demonstrated to be linear within the time frame of the assay. However, this is not the only criterion that determines initial rate and steady-state kinetics. Linearity has been observed in many non-steady-state systems (Allison and Purich, 1979).

The general rule of thumb is that initial rate conditions prevail when the substrate concentration is still within 10% of its starting value. This value is clearly acceptable for reactions that are thermodynamically quite favorable and that do not exhibit any significant product inhibition. However, if the equilibrium constant for the reaction is not large, an estimate of the degree of substrate-to-product conversion should be made. With smaller equilibrium constants, the initial-rate assay protocol necessarily must be more sensitive. In addition, if the product has very tight affinity for the enzyme, product accumulation can lead to significant errors in initial-rate determinations even in those cases where the degree of conversion is low. A preliminary test to assess this possibility is to look at initial rates in the absence and presence of a known amount of product. Where product inhibition is a problem, product can be depleted if an auxiliary enzyme system for doing this is present. Nevertheless, product inhibition can be an important regulatory scheme for an enzyme. Any kinetic characterization of an enzyme should include accurate measurements of the dissociation constants for all products. Product-inhibition studies can also assist the investigator in establishing mechanistic details of the protein undergoing evaluation (Rudolph, 1979).

In addition to establishing that the progress curve is truly linear, demonstrating that the degree of substrate depletion in the initial time course of the assay is low, and assessing the degree of product inhibition that may be present, the investigator should also demonstrate in the early stages of the kinetic studies that the rates observed are linear with respect to the total enzyme concentration ($[E]_{total}$). If true steady-state conditions are present, a plot of initial rates versus $[E]_{total}$ (in which the enzyme concentration is varied above and below the final value to be used in the initial-rate assays) should be linear.

Substrate Concentration Ranges

A wide range of substrate concentrations should be tested in preliminary studies. Such trials will provide an estimate of the Michaelis constant (K_m) and an assessment of whether substrate inhibition occurs at elevated substrate levels. These tests will also detect the possible existence of cooperativity. Once an estimate of the Michaelis constant has been made, a more detailed kinetic study can be made, in which the substrate concentration is typically varied from ~0.2 to 5.0 times the value of the apparent constant. It should be emphasized that reasonable estimates of the K_m value can only be made when initial rates have been obtained from substrate concentrations both below and above the true K_m value. For a one-substrate enzyme-catalyzed reaction:

$$K_m = [A](V_{max} - v)/v$$

where [A] is the substrate concentration, V_{max} is the maximal velocity of the reaction, and v is the velocity of the reaction at the substrate concentration. A series of substrate concentrations with their respective initial rate velocities

POLYTECHNIC UNIVERSITY LIBRARY

will provide an estimate of the Michaelis constant. In a standard double-reciprocal plot, the horizontal intercept is numerically equal to $-1/K_m$. For multisubstrate enzymes (see below), estimates of the Michaelis constants will be obtained from slope and intercept replots of the initial rate data. In addition, appropriate statistical treatment of the raw data should be provided (Wilkinson, 1961; Cleland, 1979a) for both K_m and V_{max} values (or k_{cat} values, equal to $V_{max}/[E]_{total}$, if those are reported). Note that, regardless of whether V_{max} or k_{cat} values are reported, the total protein concentration, as well as the specific activity of the enzyme used, should be reported whenever possible.

With enzyme systems utilizing two substrates, the concentrations of both substrates must be varied as described above. Typically, three solutions are prepared. One solution contains one of the substrates at its highest level. A second solution contains the second substrate at its highest level. The third solution contains the buffer, all cofactors, metal ions, auxiliary enzymes, and any other effectors, but no substrates. This third solution is used to prepare serial dilutions of the other two stock solutions. Thus, a matrix of initial rate measurements on both substrates can be evaluated from a single experimental trial. Figure 3.5.2 illustrates a typical double-reciprocal plot for a two-substrate enzyme-catalyzed reaction (where the the two substrates are A and B). In this example, Michaelis constants for the two substrates can be obtained by constructing slope and/or intercept replots of the data. For example, plotting the slopes of the five lines as a function of the reciprocal of the concentration of B will yield a straight line with a horizontal intercept numerically equal to $-1/K_B$, where K_B is the Michaelis constant for B.

For enzymes that utilize three or more substrates, the protocol described above can become unwieldy. Two procedures have been described to address the design of kinetic studies for these proteins. In the first method (Frieden, 1959), one of the three substrates is held constant at a concentration above its respective Michaelis constant that is still subsaturating. The other two are then varied as described above. This same procedure is then followed with each of the remaining substrates. Thus, multisubstrate enzymes are reduced to pseudo-two-substrate systems. One distinct advantage of this procedure is that standard graphical presentations of the data thus produced (e.g., via double-reciprocal plots) are identical to that shown in Figure 3.5.2. However, the concentration value for the third substrate have to be provided. In the second procedure, the concentration of one substrate is varied while the other substrates are held at a

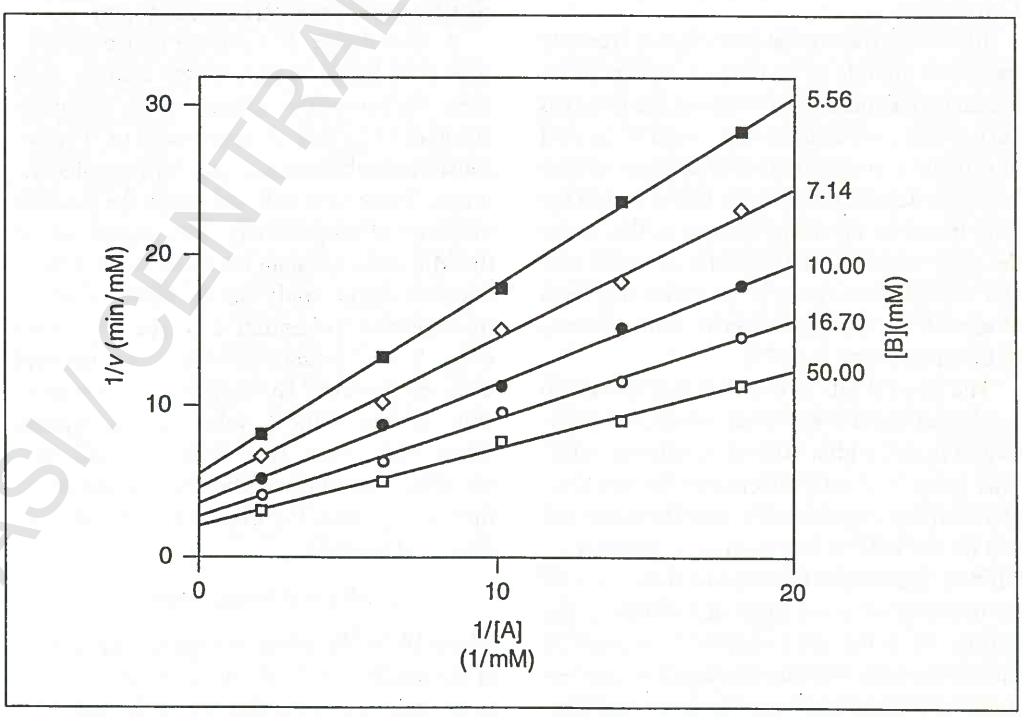


Figure 3.5.2 Double-reciprocal plot of a two-substrate enzyme-catalyzed reaction, where v = the velocity of the reaction and [A] and [B] represent the concentrations of substrates A and B respectively.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

constant ratio, although the absolute amounts are varied. This is then repeated for all remaining substrates until the concentration of each of the substrates has been varied (Fromm, 1967). One limitation with this approach is the difficulty in obtaining true Michaelis constants.

The presence of substrate inhibition can be an annoyance to an investigator who is attempting to determine the kinetic parameters of an enzyme as well as identify the binding pattern. However, the pattern of such inhibition can serve as a tool in understanding the kinetic and possibly the regulatory scheme for the enzyme. Substrate inhibition can occur as a result of a number of different causes—i.e., the formation of a dead-end complex with a particular enzyme form, the appearance of a different binding mechanism at elevated levels of substrate, the presence of an allosteric site for the substrate, or nonspecific inhibition that is due to an elevated ionic strength or another factor. When substrate inhibition occurs, in the case of multisubstrate enzymes, estimates of the K_i value (i.e., the dissociation constant for inhibition) for the substrate can be obtained from slope and intercept replots of the kinetic data. For enzymes having a single substrate, the data can be fit to an enzyme-rate expression containing a term for substrate inhibition ($v = V_{max}[S]/\{K_m + [S] + [S]^2/K_i\}$). In addition, estimates of K_m, K_i, and V_{max} can be obtained from double-reciprocal plots (Cleland, 1979b). Note that in such cases, double-reciprocal plots will be nonlinear, with the degree of curvature dependent on the magnitude of the K_i value. In these cases, it may be necessary to use a computer program to fit the data. Cleland (1979a) has provided a program to address substrate inhibition. In addition, one might consider utilizing a Marmasse plot (Marmasse, 1963). Although rarely used, this plot provides good estimates of K_m and K_i. A Marmasse plot is a plot of $1/v$ versus $(\alpha + 1/\alpha)$, where $\alpha = [A]/[A]_m$ and $[A]_m$ represents the concentration of substrate at which the velocity is highest (note that this is not V_{max}, as substrate inhibition is present.) This plot, when combined with the observation that $[A]_m = (K_m/K_i)^{1/2}$, will provide values for the kinetic parameters (Cleland, 1979b).

CONTINUOUS VERSUS STOP-TIME ASSAYS

As the best estimate for the initial rate of an enzyme-catalyzed reaction derives from the slope of the steady-state region of the progress curve, continuous assays are greatly preferable to single-point assays. If it is not possible to use a continuous assay, the investigator must demonstrate that the single-point assay is a true measure of the steady-state velocity of the reaction. In the single-point method, an aliquot of the reaction mixture is removed from the sample at a given time point after the reaction has been initiated (after a "blank" aliquot has been removed from the reaction mixture prior to the initiation step). Aliquots from at least four different time points are removed from the reaction solution; then the reaction is terminated and the amount of product formed (or substrate depleted) is determined. It should be ascertained that the termination procedure truly quenches the reaction. Once the steady-state relationship has been established, replicate kinetic analyses at each substrate concentration will provide sufficient data for a statistical analysis. The number of replicate samples that are assayed should also be reported.

COUPLED ENZYME ASSAYS

In many instances, continuous assays are made possible by the presence of auxiliary enzymes. Product formation is measured by having the product react further with another enzyme (or enzymes). Although such assay systems are convenient, precautions must be exercised to ensure that the initial rates observed reflect the enzyme that is undergoing investigation and not the auxiliary system. If the auxiliary system is influencing the initial-rate results, the reported values for the kinetic parameters will be in error and the double-reciprocals may exhibit considerable nonlinearity.

The product or products formed by the enzyme undergoing study clearly direct the choice of auxiliary enzyme(s). This coupling system should provide turnover in excess over that due to the enzyme being investigated. However, the degree of excess will vary depending upon the assay method, the nature of the coupling system, and the magnitude of the kinetic parameters for each auxiliary enzyme as measured under the same conditions being observed with the target enzyme. Once this information is in hand, the amount of auxiliary enzyme needed can be calculated (McClure, 1969; Storer and Cornish-Bowden, 1974). Test runs should then be undertaken using both the calculated and higher levels of auxiliary enzymes to insure that identical rates are obtained, indicating that a sufficient excess has been achieved. The auxiliary enzyme system should also have minimal effect on cofactors and effectors required by the target enzyme.

RUTGERS UNIVERSITY LIBRARY

Frequently, a small amount of product may be present in the reaction mixture as a contaminant. For example, commercial preparations of ATP often include a small amount of ADP. The contaminating product can be removed prior to initiation of the assay by incubating the reaction mixture with the auxiliary system. The system should also be checked to insure that no activity is observed in the absence of the auxiliary enzymes. Finally, care should be exercised when using commercial preparations for the auxiliary enzymes. Often these sources of proteins contain contaminating activities that may cause serious difficulties unless they are addressed.

BINDING STUDIES

Measurements of equilibrium ligand binding can frequently add to the characterization of an enzyme's binding mechanism, and also act as a tool in such procedures as molecular-receptor binding studies and radioimmunoassays. A wide variety of methods have been utilized in characterizing binding phenomena—e.g., gel filtration and equilibrium dialysis. The theory underlying this valuable tool can be found in any physical biochemistry text. Commonly, the binding data are analyzed with Scatchard plots (Scatchard, 1949) in which the number of moles of ligand bound per mole of total protein (ν) are plotted versus the ratio of ν to the free molar concentration of the ligand. These plots are frequently used to obtain values for the ligand dissociation constant, to identify interacting sites, and to estimate the binding capacity of the protein. Care should be exercised in this last use of the binding data, as it is easy to obtain a poor estimate of the stoichiometry of binding from Scatchard plots (Klotz, 1982, 1983). For example, a Scatchard plot of the binding of diazepam to benzodiazepine receptors is reported to provide an incorrect total number of receptor sites. Often, such problems are due to having only a few data points in the Scatchard plot at which the ligand concentration is above the half-maximum binding point. Just as in kinetic data, concentrations of the binding ligand both above and below this half-maximum point should be investigated (Klotz, 1982). Feldman (1983) has recently presented a statistical analysis associated with ligand binding data and Scatchard plots. As this paper points out, estimates of molar binding capacity should always be reported with the appropriate degree of statistical certainty. Without such statistical analyses, reports of the number of binding sites can be meaningless.

PRESENTATION OF INITIAL-RATE DATA

The description of the assay protocols used in connection with enzyme purifications, in clinical studies, and in all kinetic investigations should include a clear definition of the unit of enzyme activity. When possible, velocities should be expressed in terms of molarity changes with time, as molarity is an intensive variable. Other units, proportional to molarity, are in common use. In all cases, the conditions of the assay should be fully described, including the volume of the initiating aliquot, the specific activity, and the amount of enzyme used in each experiment. The international unit (U), defined as the conversion of 1 μmol of substrate to product per minute, is in common use. Another unit of catalytic activity, the katal, corresponding to the transformation of one mole of substrate per second, is also in use. It has been suggested that the latter unit be restricted to catalytic activities measured in a clinical context (Dybkaer, 1979).

Initial-rate data is frequently best presented via graphical displays. These visual linearizations of kinetic information provide a straightforward evaluation of the binding sequence and mechanism, as well as a ready appraisal of the kinetic parameters, reversible inhibition, cooperativity (both allosteric and hysteretic), and many other steady-state parameters. Whenever possible, slope and/or intercept replots should also be provided in the graphical display, perhaps as insets of the initial rate data. There are three methods frequently employed for depicting kinetic data, with the double-reciprocal plot (i.e., $1/v$ versus $1/[A]$; see Fig. 3.5.2) being the most common. The other common plots are $[A]/v$ versus $[A]$ and v versus $v/[A]$. Wilkinson (1961) has discussed each of these plots in terms of statistics. When the data are unweighted, the three graphical presentations are not statistically equivalent. However, this problem is considerably mitigated by proper analysis using computer methods and data weighting (Siano et al., 1975). As mentioned earlier, Cleland (1979a) has provided a number of computer programs for the analysis of data. Eisenthal and Cornish-Bowden (1974) have also introduced a novel method for graphical treatment of data that provides a useful tool for evaluating experimental error. An Eisenthal–Cornish-Bowden plot is based on the equation $(V_{max}/v) - (K_m/[A]) = 1$. Instead of a particular set of data (i.e., a single initial rate velocity with its respective substrate concentration) being represented by a point in a Car-

tesian coordinate system, the velocity (v) is denoted along the vertical axis and the substrate concentration ([A]) is denoted along the horizontal axis. The two points thus plotted are then connected by a line. Thus, whereas in the other graphical procedures the velocity and substrate-concentration data are represented by points, in the Eisenthal–Cornish-Bowden plot the same data are represented by lines. These lines, obtained from a set of initial-rate experiments, will intersect about a point equivalent to $v = V_{max}$ and $[A] = K_m$. Because median values of intersection points are used, outlier (i.e., aberrant) data will have minimal effects on estimates of the kinetic parameters. In addition, this plot is very useful in that calculations are not required and data do not have to be weighted. However, it does have limitations (Cornish-Bowden and Endrenyi, 1981, 1986). For example, the plot can become quite unwieldly with multisubstrate enzymes and in inhibition studies.

It has recently been suggested (Brand and Johnson, 1992; Johnson and Brand, 1994) that all of the graphical methods described above may be unnecessary now that personal computers and associated new methods of data analysis are available. For example, global analysis of biochemical data is a powerful tool that may prove to be quite useful in the analysis of enzyme kinetics (Beechem, 1992). However, these numerical methods have not yet been commonly applied to studies such as those of multisubstrate systems, equilibrium exchanges, or product inhibition. At this time, with proper statistical weighting of the data, multisubstrate systems are more readily appraised by one of the common graphical procedures. More importantly, these visual depictions of initial-rate data are a clear aid to any individual reviewing the kinetic experiments.

CONCLUDING REMARKS

Initial-rate assays are extremely useful tools in the detailed characterization of an enzyme. The behavior, mechanism, and regulatory properties of an enzyme can be readily assessed by straightforward assays. Such assays can frequently be completed within a day, once preliminary runs have identified the magnitude of the kinetic parameters and revealed those factors that affect activity. However, the details of the kinetic mechanism and regulatory properties of the enzyme can easily be obscured by poor assay design. The structural/functional role of an enzyme, as well as its regulation and control, can only be assessed by a reliable initial-rate assay. In addition, the data assembled from these rigorous studies permit the development of more reliable and sensitive assay protocols for protein purification and identification.

LITERATURE CITED

Allison, R.D. 1985. γ-Glutamyl transpeptidase: Kinetics and mechanism. *Methods Enzymol.* 113:419-437.

Allison, R.D. and Meister, A. 1981. Evidence that transpeptidation is a significant function of γ-glutamyl transpeptidase. *J. Biol. Chem.* 256:2988-2992.

Allison, R.D. and Purich, D.L. 1979. Practical considerations in the design of initial velocity enzyme rate assays. *Methods Enzymol.* 63:3-22.

Beechem, J.M. 1992. Global analysis of biochemical and biophysical data. *Methods Enzymol.* 210:37-54.

Bergmeyer, H.U. (ed.) 1983. Methods of Enzymatic Analysis, 3rd ed., Vols. 1-4. Verlag Chemie, Deerfield Beach, Florida.

Boehlein, S.K., Richards, N.G.J., and Schuster, S.M. 1994. Glutamine-dependent nitrogen transfer in *Escherichia coli* asparagine synthetase B. *J. Biol. Chem.* 269:7450-7457.

Bowen, W.J. and Kerwin, T.P. 1955. The purification of myokinase with an ion exchange resin. *Arch. Biochem. Biophys.* 57:522-524.

Brand, L. and Johnson, M.L. (eds.). 1992. Numerical computer methods. *Methods Enzymol.* Vol. 210.

Cleland, W.W. 1979a. Statistical analysis of enzyme kinetic data. *Methods Enzymol.* 63:103-138.

Cleland, W.W. 1979b. Substrate inhibition. *Methods Enzymol.* 63:500-513.

Cornish-Bowden, A. and Endrenyi, L. 1981. Fitting of enzyme kinetic data without prior knowledge of weights. *Biochem. J.* 193:1005-1008.

Cornish-Bowden, A, and Endrenyi, L. 1986. Robust regression of enzyme kinetic data. *Biochem. J.* 234:21-29.

Deans, N.L., Allison, R.D., and Purich, D.L. 1992. Steady-state kinetic mechanism of bovine brain tubulin:tyrosine ligase. *Biochem. J.* 286:243-251.

Dybkaer, R.. 1979. International Union of Pure and Applied Chemistry and International Federation of Clinical Chemistry. IUPAC Section of Clinical Chemistry; Commission on Quantities and Units in Clinical Chemistry and IFCC Committee on Standards, Expert Panel on Quantities and Units; Approved Recommendation (1978) Quantities and units in clinical chemistry. *Clin. Chim. Acta* 96:157F-183F.

Eisenthal, R. and Cornish-Bowden, A. 1974. The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochem. J.* 139:715-720.

CHILE

BOTANICAL GARDEN LIBRARY

Feldman, H.A. 1983. Statistical limits in Scatchard analysis. *J. Biol. Chem.* 258:12865-12867.

Fierke, C.A. and Hammes, G. 1995. Transient kinetic approaches to enzyme mechanisms. *Methods Enzymol.* 249:1-32.

Frieden, C. 1959. Glutamate dehydrogenase III. The order of substrate addition in the enzymatic reaction. *J. Biol. Chem.* 234:2391-2396.

Fromm, H.J. 1967. The use of competitive inhibitors in studying the mechanism of action of some enzyme systems utilizing three substrates. *Biochim. Biophys. Acta* 139:221-230.

Johnson, M.L. and Brand, L. (eds.). 1994. Numerical computer methods, part B. *Methods Enzymol.* 240:xi-xiii, 1-36, 51-170, 181-198, 311-322, 781-816.

Klotz, I.M. 1982. Numbers of receptor sites from Scatchard graphs: Facts and fantasies. *Science* 217:1247-1249.

Klotz, I.M. 1983. Number of receptor sites from Scatchard and Klotz graphs: A constructive critique. *Science* 220:981.

Klump, H., Di Ruggiero, J., Kessel, M., Park, J.-B., Adams, M.W.W., and Robb, F.T. 1992. Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. *J. Biol. Chem.* 267:22681-22685.

Lienhard, G.E. and Secemski, I.I. 1973. P^1, P^5-Di(adenosine-5′)pentaphosphate, a potent multisubstrate inhibitor of adenylate kinase. *J. Biol. Chem.* 248:1121-1123.

Marmasse, C. 1963. Enzyme inhibition by excess substrate. *Biochim. Biophys. Acta* 77:530-535.

McClure, W.R. 1969. A kinetic analysis of coupled enzyme assays. *Biochem.* 8:2782-2786.

Nordlie, R.C. 1982. Kinetic examination of enzyme mechanisms involving branched reaction pathways: A detailed consideration of multifunctional glucose-6-phosphatase. *Methods Enzymol.* 87:319-353.

O'Sullivan, W.J. and Smithers, G.W. 1979. Stability constants for biologically important metal-ligand complexes. *Methods Enzymol.* 63:294-336.

Purich, D.L. and Fromm, H.J. 1972. Inhibition of rabbit muscle adenylate kinase by the transition state analogue, P^1,P^4-di(adenosine-5′)tetraphosphate. *Biochim. Biophys. Acta* 276:563-567.

Rudolph, F.B. 1979. Product inhibition and abortive complex formation. *Methods Enzymol.* 63: 411-436.

Scatchard, G. 1949. The attraction of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.* 51:660-673.

Seelig, G.F. and Meister, A. 1985. Glutathione biosynthesis: γ-Glutamylcysteine synthetase from rat kidney. *Methods Enzymol.* 113:379-390.

Siano, D.B., Zyskind, J.W., and Fromm, H.J. 1975. A computer program for fitting and statistically analyzing initial rate data applied to bovine hexokinase type III isozyme. *Arch. Biochem. Biophys.* 170:587-600.

Silverstein, E. and Sulebele, G. 1969. Catalytic mechanism of pig heart mitochondrial malate dehydrogenase studied by kinetics at equilibrium. *Biochemistry* 8:2543-2550.

Storer, A.C. and Cornish-Bowden, A. 1974. The kinetics of coupled enzyme reactions. *Biochem. J.* 141:205-209.

Tate, S.S. and Meister, A. 1978. Serine-borate complex as a transition-state inhibitor of γ-glutamyl transpeptidase. *Proc. Natl. Acad. Sci. U.S.A.* 75:4806-4809.

Wilkinson, G.N. 1961. Statistical estimations in enzyme kinetics. *Biochem. J.* 80:324-332.

Williams, J.W., and Morrison, J.F. 1979. The kinetics of reversible tight-binding inhibition. *Methods Enzymol.* 63:437-467.

KEY REFERENCES

Bergmeyer, 1983. See above.

Detailed collection of enzyme assay protocols and methods for a wide variety of enzymes, including descriptions of many special techniques: e.g., fluorometry, turbidimetry, luminometry, and microtechniques.

Dixon, M. and Webb, E.C. 1979. Enzymes, 3rd ed. Academic Press, New York.

Classic text on the basics of enzymology; Chapter 4 deals specifically with enzyme kinetics.

Fromm, H.J. 1975. Initial Rate Enzyme Kinetics. Springer-Verlag, New York.

Valuable introductory text with considerable details on experimental design.

Kuby, S.A. 1990. A Study of Enzymes, vol. I. CRC Press, Boca Raton, Florida

Detailed text on steady-state kinetics. Contains useful chapters on the effects of metal cofactors and transient phase kinetics.

Purich, D.L. (ed.) 1979. Enzyme kinetics and mechanism, part A. *Methods Enzymol.* Vol. 63.

Describes a variety of initial rate methods and inhibitor studies for characterizing enzyme-catalyzed reactions.

Purich, D.L. (ed.) 1980. Enzyme kinetics and mechanism, part B. *Methods Enzymol.* Vol. 64.

Describes a number of isotopic probes for studying enzyme mechanisms and contains valuable chapters on allosterism, hysteresis, immobilized systems, and processivity.

Purich, D. L. (ed.) 1982. Enzyme kinetics and mechanism, part C. *Methods Enzymol.* Vol. 87.

Describes methods for characterizing intermediates in enzyme-catalyzed reactions and using stereochemical probes for enzyme mechanisms; includes additional initial-rate and inhibitor methods and discusses further uses of isotopic probes.

BCU IAS SETTAI UNIVERSITY LIBRARY

Purich, D.L. (ed.) 1995. Enzyme kinetics and mechanism, part D. *Methods in Enzymol.* 249:3-662.

Covers in detail specialized topics in enzyme kinetics including transition-state approaches, kinetic probes with site-directed mutagenesis, partition analysis, positional isotope exchange, interfacial catalysis, and hydrogen tunneling.

Contributed by R. Donald Allison
University of Florida
Gainesville, Florida

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Biotinylation of Proteins in Solution and on Cell Surfaces

A protein is typically labeled with the water-soluble vitamin biotin to tag it for later recovery or for detection with biotin-binding proteins, such as avidin or streptavidin. Applications of this approach include identification and purification of cell surface–exposed proteins, measurement of ligand internalization kinetics, and temporal changes in protein distributions inside cells using light or electron microscopy. Many modified biotins, capable of reacting with protein lysyls, cysteinyls, or carbohydrates, are commercially available (see Table 3.6.1). It is usually possible to find at least one biotin reagent that can derivatize a given protein without affecting the function of interest. For instance, most enzyme activities are independent of protein modification on carbohydrate residues. Also, when used with enzyme-linked or radiolabeled avidin or streptavidin, biotinylation can be an effective technique for visualizing proteins that do not bind well to dyes after separation by SDS–polyacrylamide gel electrophoresis (SDS-PAGE; *UNIT 10.1*).

This unit contains three protocols for biotinylation of isolated proteins: attaching biotin to primary amines (e.g., ε-amino groups of lysyl residues; see Basic Protocol 1); attaching biotin to sulfhydryls (i.e., thiol groups of cysteinyl residues; see Basic Protocol 3); and attaching biotin to carbohydrate residues on proteins (see Basic Protocol 4). As biotinylation of lysyl and cysteinyl residues may alter protein function, modification of protein carbohydrates, which is usually innocuous, may be preferable if intact protein function is required (e.g., for activity assays or affinity purification). Biotinylation of cysteinyl thiols requires that disulfide bonds in isolated proteins be reduced before labeling (see Support Protocol 1).

Mild reaction conditions (e.g., in aqueous buffers at pH 7.5 to 8.5) and the small size of the biotin moiety also make biotinylation ideal for labeling surface proteins on living cells (see Basic Protocol 2). Such site-specific labeling enables investigators to monitor and/or effect subcellular purifications. Finally, this unit includes a brief description of methods for detecting biotinylated proteins (see Support Protocol 2).

COVALENT ATTACHMENT OF BIOTIN TO LYSINES

BASIC PROTOCOL 1

This protocol is a specialized version of the succinimidyl ester amidation reaction described in Basic Protocol 1 of *UNIT 15.2*. The reagent used, NHS-LC-biotin (see Fig. 3.6.1), has an *N*-hydroxysuccinimide (NHS) ester group to increase reactivity with primary amines and a long-chain (LC) spacer arm to extend the biotin from the protein surface, thus improving binding of subsequent reagents (e.g., avidin). The protein of interest is incubated with NHS-LC-biotin. Glycine is added to stop the reaction, and biotinylated proteins are purified by dialysis. This protocol works well over the range of 50 µl to 50 ml (and possibly for much larger volumes), so the volume of the reaction mixture should be chosen to suit the investigator's needs.

If biotinylation of lysines results in loss of protein function, biotinylation of thiols (see Basic Protocol 3) or carbohydrate groups (see Basic Protocol 4) may prove useful.

NOTE: NHS-biotin should be stored in the presence of desiccant at –20°C. When the reagent is thawed, it should be warmed thoroughly to room temperature before the container is opened.

Contributed by Elizabeth J. Luna
Current Protocols in Protein Science (1996) 3.6.1-3.6.15
Copyright © 1996 by John Wiley & Sons, Inc.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

Table 3.6.1 Commercially Available Biotinylating Reagents

Reagent	Supplier[a]	Molecular weight	Comments
Specific for reaction with primary amines (see Basic Protocols 1 and 2, Support Protocol 2)			
NHS-biotin	Pierce, Sigma, Vector Labs	341.4	Soluble in DMF or DMSO; store all NHS reagents at –20°C in presence of desiccant
Sulfo-NHS-biotin	Pierce, Sigma, Vector Labs	443.4	Water soluble
NHS-LC-biotin	Pierce, Sigma, Vector Labs	454.5	Soluble in DMF or DMSO; spacer arm improves accessibility of protein-bound biotin to affinity matrices
NHS-SS-biotin	Pierce	606.7	Soluble in DMF; cleavable by disulfide reduction
2-Iminobiotin-NHS	Calbiochem, Pierce	421.3	Soluble in DMF; exhibits pH-sensitive binding to avidin
Specific for reaction with sulfhydryls (see Basic Protocol 3, Support Protocols 1 and 2)			
Iodoacetyl-LC-biotin	Pierce	510.4	Soluble in DMF or DMSO; reacts with amino groups above pH 7.5
Biotin-LC maleimide	Sigma	479.6	Soluble in DMF; reacts with amino groups above pH 7.5
Maleimido-propionyl biocytin	Molecular Probes, Pierce	523.6	Soluble in DMF; reacts with amino groups above pH 7.5
Biotin-HPDP	Pierce	539.8	Soluble in DMSO; cleavable by disulfide reduction, regenerating initial cysteinyl
Specific for reaction with (oxidized) carbohydrates (see Basic Protocol 4, Support Protocol 2)			
Biotin-LC-hydrazide	Calbiochem, Pierce, Vector Labs	258.3	Soluble in water (to ~5 mM) and DMSO (to ~50 mM)
Biotin-hydrazide	Molecular Probes, Pierce, Sigma	371.5	Soluble in water (to ~5 mM) and DMSO (to ~50 mM)

[a]See *SUPPLIERS APPENDIX* for addresses and phone numbers of manufacturers.

Materials

~1 mg/ml purified protein
Coupling buffer (see recipe), ice cold
2 mg/ml NHS-LC-biotin (Pierce, Vector Labs) in *N,N*-dimethylformamide (DMF), room temperature, prepared just before use
50 mM glycine in coupling buffer
0.22-µm filter: e.g., Spin-X centrifuge filter (Costar) or syringe filter (Nalgene)
Additional reagents and equipment for dialysis (*UNIT 4.4* and *APPENDIX 3B*)

1. If necessary, dialyze the purified protein into coupling buffer to remove traces of interfering buffer components that may have been introduced during purification.

Complete exchange of buffer components can be assumed after a 48-hr dialysis against 100 vol of coupling buffer with three to four buffer changes. However, if the initial concentration of the interfering agent is relatively low, a 16- to 24-hr dialysis with two buffer changes is generally sufficient.

Azide, glycine, ammonium ions, and Tris are common buffer components that interfere with the biotin labeling reaction. Although these reagents should be avoided, reagents without primary or secondary amino groups can be included to stabilize the protein or to prevent labeling at functional sites. For instance, NaCl, $CaCl_2$, sucrose, glycerol, Triton X-100, and HEPES all can be used in the coupling buffer, if required.

CSU IAS / GENERAL UNIVERSITY LIBRARY

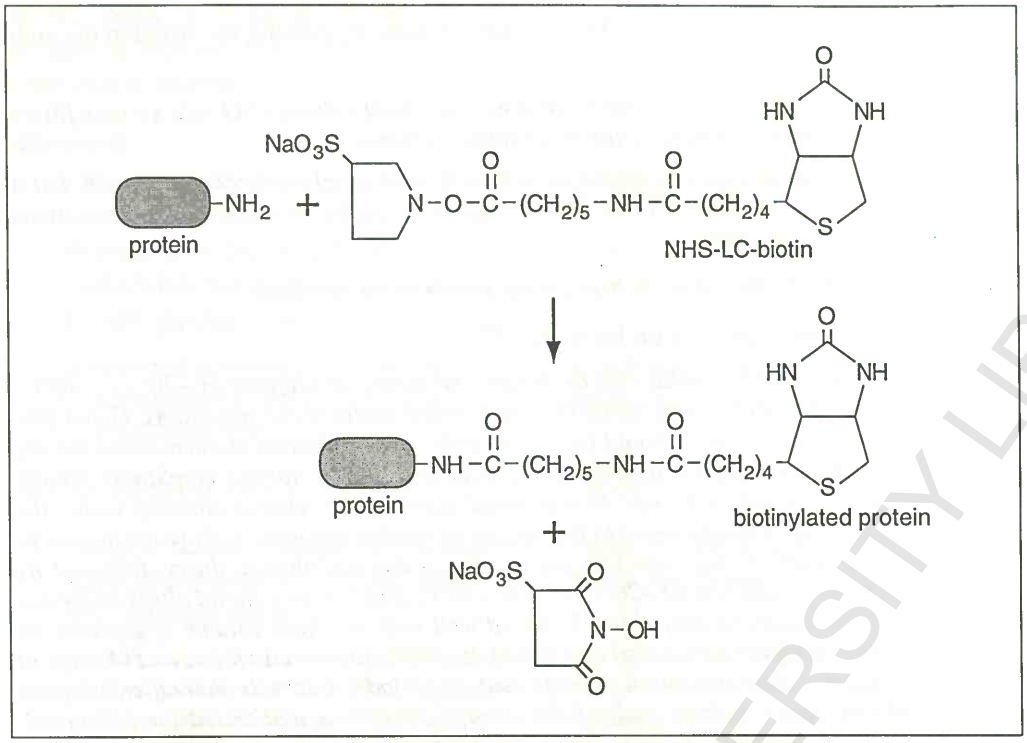


Figure 3.6.1 Reaction of *N*-hydroxysuccinimide (NHS)-LC-biotin with protein.

2. Rapidly add the freshly prepared suspension of NHS-LC-biotin in DMF to the protein in coupling buffer, using 1 vol biotin in DMF to 20 vol protein solution.

   The biotin/protein mixture may become turbid because NHS-LC-biotin forms a fine suspension in water, rather than a true solution. If the protein is sensitive to organic solvents or if the highest affinity binding to avidin afforded by the LC spacer is not necessary, sulfo-NHS-biotin (Pierce, Sigma, or Vector Labs) can be substituted for NHS-LC-biotin. Sulfo-NHS-biotin is very water soluble and can be added at the same final biotin/protein molar ratio (~10:1) from a freshly prepared aqueous stock solution without precipitate formation.

3. Incubate ~1 hr on ice. Mix occasionally.

   Optimal extents of biotinylation must be determined empirically for each protein and are a compromise between high levels of incorporated biotin and retention of protein function. The biotin/protein ratio during labeling and the length of labeling time recommended here have worked well for a number of proteins. However, biotin-to-protein ratios as low as 1:1 and as high as 60:1 have been reported. Similarly, optimal labeling times may vary between 1 and 18 hr.

4. Add 1 vol of 50 mM glycine to bind unreacted biotin.

5. Remove glycine and free biotin by exhaustive dialysis against coupling buffer, or another buffer of choice. Dialysis for ≥36 hr against three or more changes of 500 vol each is recommended.

   Once the labeling step is complete, any buffer, including Tris·HCl, can be used for dialysis (UNIT 4.4 & APPENDIX 3B). Unincorporated sulfo-NHS-biotin can be removed quickly by gel-filtration on an Econo-Pac 10DG desalting column (Bio-Rad) or a Sephadex G-25 column (Pharmacia Biotech; see UNIT 8.3). However, because any suspended micelles of unreacted NHS-LC-biotin will coelute in the void volume with the protein, proteins labeled with NHS-LC-biotin should be dialyzed overnight against one change of buffer before gel filtration.

BOULDER LIBRARY UNIVERSITY

6. For long-term storage, sterilize biotinylated proteins by passing the solution through a 0.22-µm filter.

   Spin-X centrifuge filters (Costar) work well for small volumes (<1 ml); syringe filters (Nalgene) are recommended for much larger volumes.

   Sodium azide also may be added up to 0.02% (w/v) to inhibit bacterial growth, but it is most effective below pH 7.0. Thimerosal (0.1% [w/v]) is more useful in solutions with higher pH values.

   CAUTION: *Sodium azide is toxic; wear gloves when handling.*

7. Store biotinylated proteins on ice or at 4°C.

   Some biotinylated proteins can be frozen and stored in aliquots at –20° or –70°C; others are reported to aggregate or lose function under these conditions. Generally, the biotinylated protein should be stored under the conditions recommended for the unlabeled protein. If no information is available about storage conditions for the protein of interest, it is wise to test small aliquots for relative stability under the following sets of conditions: (1) fast frozen in sterile cryogenic vials (Corning) or in test tubes with O-ring caps (Sarstedt) using a dry ice/ethanol slurry, followed by storage of aliquots (1a) at –20°C, (1b) at –70°C, and (1c) in a liquid nitrogen Dewar flask. Other aliquots should be (2) mixed well with an equal volume of glycerol (to make a 50% glycerol solution) and stored at –20°C, (3) stored on ice, and (4) kept at 4°C. Shortly after the initial storage and at periodic intervals thereafter, aliquots should be removed from each of the storage conditions and tested for biological, binding, or enzymatic activity. Once the most efficient storage condition has been determined, future preparations of the biotinylated protein should be stored accordingly in aliquots sufficient for one, or a few, experiments. Repeated cycles of freezing and thawing should be avoided.

   CAUTION: *Tubes stored under liquid nitrogen can explode on warming if liquid nitrogen has gotten inside the tube. Loosen caps on tubes immediately after taking them out of a liquid nitrogen Dewar flask.*

   The concentration of the biotinylated protein can be determined using any of the assays described in UNIT 3.4, and the concentration of protein-associated biotin can be measured in a variety of solid-phase competition assays with labeled avidin and known amounts of free biotin (Savage et al., 1992). However, because most assays with biotinylated protein are either qualitative or comparative, many investigators merely verify that biotinylation has occurred, i.e., by analysis of electrotransferred protein after SDS-PAGE (UNIT 10.10).

BASIC PROTOCOL 2

SIDE-SPECIFIC BIOTINYLATION OF PROTEINS ON CELL SURFACES

Biotinylation of proteins at the extracellular surface of living cells is a fast, nonradioactive means of labeling and tracking the plasma membrane during cell fractionation (Ingalls et al., 1986; UNIT 4.2). The biotinylated proteins also are easily isolated after membrane solubilization and affinity chromatography on columns containing bound avidin or streptavidin (Chia and Luna, 1989).

Biotinylation of living cells requires the water-soluble biotin derivative sulfo-NHS-biotin. The starting material is intact cells, either freshly isolated or from suspension or adherent cell cultures; control experiments starting with lysed cells or isolated membranes are, however, recommended to verify that cell integrity is maintained during the procedure. Cells are reacted with freshly dissolved sulfo-*N*-hydroxysuccinimide (NHS)-biotin for 30 min with mixing. Glycine is added to bind unreacted biotin, and cells are washed and lysed to prepare plasma membranes or solubilized proteins.

SCU IAS CENTRAL UNIVERSITY LIBRARY

NOTE: NHS-biotin should be stored in the presence of desiccant at –20°C. When the reagent is thawed, it should be warmed thoroughly to room temperature before the container is opened.

Materials

10^7 to 10^{10} cells of interest, in suspension (e.g., erythrocytes, leukocytes, or HeLa cells), or cultures of adherent cells
Physiological buffer, pH 7.4 (see recipe), ice cold
Sulfo-NHS-biotin (Pierce or Vector Labs)
10 mM glycine in physiological buffer

Additional reagents and equipment for lysing cells and isolating plasma membranes (*UNIT 4.2* and *UNIT 4.3*) or for solubilizing proteins for SDS gel electrophoresis (*UNIT 10.1*)

NOTE: Carry out all steps in this protocol at 4°C or on ice.

1. Remove serum proteins by washing the cells of interest at least three times with ice-cold physiological buffer, centrifuging cells in suspension for 5 min at ~500 × *g*, 4°C, between washes.

   Adherent cells are most easily washed and labeled while still attached to tissue culture flasks. To avoid labeling cytoplasmic proteins, the biotinylating agent should be thoroughly blocked and removed before dissociating attached cells from the flask (i.e., detach the cells after step 5). The buffer can be varied, depending on the preferences of the cells and the investigator. Dulbecco's PBS or Earle's balanced salts (e.g., Sigma) also work well. In contrast, PBS containing only ~5 mM phosphate (APPENDIX 2E) may not have sufficient buffering capacity. As in Basic Protocol 1, the principal consideration is to omit reagents containing primary amino groups that could compete for binding to the sulfo-NHS-biotin.

2. During the final cell wash, dissolve the sulfo-NHS-biotin in 0.1 vol ice-cold physiological buffer for a final concentration of 2 mg/ml (~5 mM). For instance, for ~10^{10} cells in a volume of 9 ml, dissolve 20 mg sulfo-NHS-biotin in 1 ml physiological buffer.

   If labeling attached cells, mix the sulfo-NHS-biotin with the entire amount of buffer immediately before adding the mixture to the cells. Minimize the amount of hydrolysis of the succinimide ester by limiting its exposure to water in the absence of cells.

3. Gently suspend the cells by adding ice-cold physiological buffer first. Then add 1/10 vol of dissolved sulfo-NHS-biotin. Mix gently without lysing the cells.

   Although labeling of cell-surface lysines proceeds faster as the pH of the buffer is raised from pH 7.4 to 9.0, so does the hydrolysis of the sulfo-NHS-biotin. In addition, many cells are lysed in higher-pH buffers.

4. Incubate 30 min at 4°C with gentle rotation or swirling.

   This procedure should label only surface-exposed proteins. Higher temperatures will result in increased biotinylation of membrane proteins, but much of the increase may be due to endocytic membrane turnover.

5. Block unreacted biotinylating reagent for ≥10 min by adding an equal volume of 10 mM glycine in physiological buffer. Wash cells with buffer three times as in step 1. Monitor the cells for lysis during centrifugation, using phase-contrast or differential interference contrast (DIC) light microscopy.

   If the cells are too fragile to withstand repeated washes, incubate with the glycine at 4°C for 15 to 30 min, then wash only one or two times. Blocking is faster at elevated pH, but pH 7.4 is adequate as long as the cells are not permeabilized during the washes.

BOSTON UNIVERSITY LIBRARY

6. Lyse the cells and isolate plasma membranes (see *UNIT 4.2*) or solubilize proteins for SDS-PAGE (see *UNIT 10.1*, Basic Protocol 1). Store biotinylated membranes in aliquots at –20°C.

It is usually a good idea to perform a parallel, control experiment in which cells are biotinylated after lysis or in which isolated membranes are biotinylated. The control procedure should label both surface-exposed and intracellular polypeptides, providing evidence for cellular integrity during the initial labeling procedure. The control can also demonstrate that if a protein of particular interest remains unlabeled under surface-labeling conditions but is labeled in isolated membranes or lysates, it is inherently capable of reacting with the sulfo-NHS-biotin reagent and, thus, is either a cytosolic protein or a membrane protein without accessible external lysines.

BASIC PROTOCOL 3

COVALENT ATTACHMENT OF BIOTIN TO THIOLS

An alternative to biotinylating lysyl groups (see Basic Protocol 1) is biotinylation of thiols or carbohydrate groups (see Basic Protocol 4). Several modified biotins react specifically with sulfhydryls (see Table 3.6.1). This protocol is described for iodoacetyl-LC-biotin in *N,N*-dimethylformamide (DMF). However, guidelines for using biotin-maleimide and *N*-(6-[biotinamido]hexyl)-3′-(2′-pyridyldithio)propionamide (biotin-HPDP) reagents are given in step annotations. Purified protein or membrane fractions are first subjected to disulfide reduction (see Support Protocol 1). The reduced material is mixed with iodoacetyl-LC-biotin. Unreacted components are then separated from labeled proteins by gel filtration or dialysis before the proteins are filter sterilized and stored.

Materials

0.5 mg/ml protein or cell fraction of interest, e.g., purified membranes
100 mM sodium phosphate/5 mM EDTA, pH 8.3 (*APPENDIX 2E*)
4 mM iodoacetyl-LC-biotin (Pierce) in DMF, freshly prepared or thawed from storage at –20°C

Additional reagents and equipment for reducing protein disulfides (see Support Protocol 1), dialysis (*UNIT 4.4* and *APPENDIX 3B*), gel-filtration chromatography (*UNIT 8.3*), or centrifugal filtration (*UNIT 15.1*, Support Protocol 3)

1. If necessary, reduce protein disulfides in the protein or cell fraction of interest (see Support Protocol 1) and remove the thiol-reducing agent by dialysis (*UNIT 4.4* and *APPENDIX 3B*), gel-filtration chromatography (*UNIT 8.3*), or centrifugal filtration (*UNIT 15.1*). Use the sodium phosphate/EDTA buffer to make up the solution of reduced protein.

If samples must be dialyzed for prolonged periods, nitrogen gas should be bubbled through the dialysis buffer to help exclude the atmospheric oxygen that will slowly (1 to 7 days) oxidize the reduced disulfides. EDTA chelates free metal ions that can catalyze this oxidation.

2. Add 20 µl of 4 mM iodoacetyl-LC-biotin in DMF for each milliliter of protein or cell fraction in the sodium phosphate/EDTA buffer.

For proteins with an "average" number of cysteines, i.e., 1.7 mol % (Table A.1A.1), this provides an equimolar ratio of iodoacetyl-LC-biotin. The cost and relative insolubility of biotinylated alkylating reagents contraindicate their use in a 5- to 10-fold molar excess over total thiols, as described in Basic Protocol 1.

Depending on the protein and activity to be studied, another thiol-directed biotinylating reagent may be preferable. At pH 8.3, iodoacetyl-LC-biotin reacts irreversibly and specifically with free sulfhydryls. Maleimide reagents (biotin-LC-maleimide and maleimidopropionyl biocytin) also irreversibly derivatize sulfhydryls but will react with some primary amines unless the pH is lowered to between 6.5 and 7.5. Biotin-

CUHK LIBRARIES

HPDP reacts rapidly and specifically through disulfide exchange with free protein sulfhydryls; consequently, proteins derivatized with biotin-HPDP should not be exposed to reduced thiol agents unless reversal of the labeling procedure is desired.

3. Incubate ~2 hr at room temperature or overnight at 0° to 4°C.

4. Remove unreacted reagent by gel-filtration chromatography (*UNIT 8.3*) or dialysis (*UNIT 4.4* and *APPENDIX 3B*).

5. Filter the solution through an 0.22-µm filter and store biotinylated proteins at 0° to 4°C or biotinylated membranes in aliquots at –20°C.

REDUCTION OF DISULFIDE BONDS

SUPPORT PROTOCOL 1

Reducing disulfide bonds may be necessary to create the free sulfhydryls on protein cysteines needed for biotinylation with thiol-directed biotin reagents (see Basic Protocol 3). Disulfide reduction can also be used to elute proteins tagged with the amine-directed reagent *N*-hydroxysuccinimide (NHS)-SS-biotin (see Basic Protocol 1) from avidin affinity columns. This protocol gives an effective method for reducing a thiol-sensitive disulfide bond that is not essential for the quaternary structure, e.g., that in immunoglobulins. The protein of interest is dialyzed into reducing buffer and mixed with reducing agent at elevated temperature (i.e., 37°C). Excess reductant is removed by gel-filtration chromatography (*UNIT 8.3*). The reduced protein is then ready for immediate biotinylation. In general, disulfide reduction can be assumed to be innocuous, if not advantageous, for proteins that reside in the cell cytoplasm, but it may perturb function in extracellular proteins or in the extracellular domains of transmembrane proteins.

Materials

1 to 10 mg/ml purified protein
Reducing buffer: 100 mM sodium phosphate/5 mM EDTA, pH 6.0 (*APPENDIX 2*)
Reducing agent: 50 mM dithiothreitol (DTT),100 mM 2-mercaptoethanol (2-ME), *or* 100 mM mercaptoethylamine·HCl (Sigma)
Econo-Pac 10DG desalting column (Bio-Rad) or Sephadex G-25 column (Pharmacia Biotech)
Additional reagents and equipment for dialysis (*UNIT 4.4* and *APPENDIX 3B*) and gel-filtration chromatography (*UNIT 8.3*)

1. If necessary, dialyze purified protein of interest into reducing buffer.

2. Add 1 vol of 1 to 10 mg/ml purified protein in reducing buffer to an equal volume of reducing agent. Incubate 1 hr at 37°C with occasional mixing.

   Each mole of DTT contains two thiols, so only half the concentration (50 mM) is required for the same concentration of reducing equivalents used for other reducing agents (100 mM 2-ME or 100 mM mercaptoethanolamine·HCl).

3. During the incubation in step 2, equilibrate an Econo-Pac 10DG desalting column or a Sephadex G-25 column with reducing buffer.

   The column should contain ≥10 vol of gel-filtration matrix (i.e., at least five times more volume than that of the reaction in step 2).

4. Cool the reaction to room temperature. Remove excess reducing agent by gel-filtration chromatography (*UNIT 8.3*) using the equilibrated column. Collect at least ten fractions, each containing ~1 vol (i.e., the same volume as the initial volume of protein solution).

   Alternatively, excess reducing agent may be removed by centrifugal filtration (UNIT 15.1, Support Protocol 3) or dialysis (UNIT 4.4 and APPENDIX 3B). If prolonged dialysis is

BOSTON CENTRAL UNIVERSITY LIBRARY

required, reoxidation of protein disulfides can be slowed by including 5 mM EDTA and/or by bubbling nitrogen gas through the dialysis buffer.

5. Read the absorbance of each fraction at 280 nm (A_{280}). Pool fractions that correspond to the first peak of absorbance.

6. As soon as possible after removing excess reducing agent, use the pooled fractions in cysteinyl-directed biotinylations (see Basic Protocol 3).

BASIC PROTOCOL 4

COVALENT ATTACHMENT OF BIOTIN TO CARBOHYDRATES

Biotinylation of protein carbohydrates requires oxidation of vicinal carbohydrate hydroxyl groups to aldehydes that can bind biotin-LC-hydrazide. Sodium *m*-periodate is used as the oxidizing agent for glycoproteins in solution, but enzyme-mediated oxidations can be used to oxidize glycoproteins of cells or membrane fractions. Briefly, glycoproteins are oxidized, excess oxidizing agent is removed by gel-filtration, biotin reagent is added, and the biotinylated glycoprotein is purified in a microconcentrator.

Materials

0.1 to 0.2 mg/ml glycoprotein
Labeling buffer: 100 mM sodium acetate, pH 5.5 (*APPENDIX 2E*)
30 mM and 3 mM sodium *m*-periodate (Sigma, mol. wt. 213.9) in water, freshly prepared and stored in the dark
5 mM biotin-LC-hydrazide (Calbiochem, Pierce, or Vector Labs) in labeling buffer (store ≤2 months at 4°C)
Stop solution: 100 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*)
Econo-Pac 10DG desalting column (Bio-Rad) or Sephadex G-25 column (Pharmacia Biotech)
Microconcentrator
Additional reagents and equipment for dialysis (*UNIT 4.4* and *APPENDIX 3B*) and gel-filtration chromatography (*UNIT 8.3*)

1. If needed, dialyze the glycoprotein into labeling buffer.

2. Add 1 vol of 30 mM or 3 mM sodium *m*-periodate to 2 vol of 0.1 to 0.2 mg/ml glycoprotein in labeling buffer for a final periodate concentration of either 10 mM or 1 mM. Incubate ~30 min at room temperature in the dark.

   Because proteins differ in sensitivity to periodate, it is a good idea to try at least two different concentrations of the oxidant. For antibodies, O'Shannessy et al. (1984) report excellent results with 10 mM sodium m-periodate. For labeling a transmembrane glycoprotein that requires a detergent for solubilization, avoid the use of octylglucoside, which also reacts with periodate.

   Cell-surface carbohydrates on erythrocytes and on purified erythrocyte membranes have been biotinylated by labeling cells or purified membranes for 2 hr at room temperature with a solution containing 0.03 U/ml Vibrio cholerae neuraminidase (Boehringer Mannheim), 3 U/ml Dactylium dendroides galactose oxidase, and 0.4 mM biotin-LC-hydrazide in physiological buffer (Bayer et al., 1988). With many cells and cell membranes, however, those conditions will promote cell lysis and/or proteolysis.

3. During the incubation period (step 2), equilibrate an Econo-Pac 10DG desalting column or Sephadex G-25 column with labeling buffer.

   The column should contain ≥15 vol of gel-filtration matrix (i.e., at least five times more volume than that of the reaction in step 2).

PCU IASI CENTRAL UNIVERSITY LIBRARY

4. Remove unincorporated sodium *m*-periodate by gel-filtration chromatography (*UNIT 8.3*) using the equilibrated column. Collect at least ten fractions, each containing ~1 vol (i.e., the same volume as the initial volume of sodium *m*-periodate).

5. Read the absorbance of each fraction at 280 nm (A_{280}). Pool fractions that correspond to the first peak of absorbance.

6. Add 1 vol of 5 mM biotin-LC-hydrazide in labeling buffer. Incubate 1 hr at room temperature with occasional mixing.

   The amount of dilution of protein on the gel-filtration column can vary without significantly affecting the efficiency of the reaction.

7. Terminate the reaction by adding 5 vol stop solution.

   Tris, glycine, ammonium ions, and other buffer components containing primary amines should be avoided during the oxidation and labeling steps because they form reversible Schiff bonds with the carbohydrate aldehydes. These Schiff bonds will inhibit hydrazone bond formation. Primary amines are, however, good for stopping the labeling reaction. If the biotinylated protein is expected to be less than ~0.5 mg/ml after concentration, a carrier protein, such as BSA, can be added after the stop solution. Carrier protein helps to eliminate protein loss during concentration but will complicate quantification of the percent recovery of the biotinylated glycoprotein and of the number of biotins incorporated per glycoprotein molecule.

8. Remove unreacted biotin-LC-hydrazide by centrifugation in a microconcentrator with a molecular weight cutoff (MWCO) that is less than half that of the molecular weight of the protein in solution (*UNIT 4.4*). After concentrating the protein, suspend the sample to the initial volume with labeling buffer and reconcentrate. Repeat until the buffer has been changed a total of three times and the glycoprotein has reached the desired concentration.

   Any buffer in which the glycoprotein is stable can be substituted for labeling buffer at this step. To estimate the final glycoprotein concentration, assume no more than 50% loss during the entire procedure, as long as the protein is known to be well-behaved in solution. If precision is necessary, the molar ratio of biotin to glycoprotein can be calculated after determining protein concentrations (UNIT 3.4) and biotin concentrations (e.g., by solid-phase binding assays with avidin; Savage et al., 1992), after taking into account any added carrier protein. However, the estimate of 50% loss during experimental manipulation generally provides a reasonable starting point for most qualitative experiments (and is seldom off by more than a factor of 2).

9. Store the concentrated glycoprotein in aliquots at –70° or –20°C.

   For discussion of storage conditions, see Basic Protocol 1 (step 7 annotation).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

DETECTION OF BIOTINYLATED PROTEINS

Once proteins or cells are biotinylated (see Basic Protocols 1 to 4), the next step is often detection (e.g., in the course of performing fluorescence or electron microscopy, assaying enzyme activity, or sizing proteins). Experimental samples and biotinylated standards can be analyzed simultaneously. This protocol gives brief instructions for separating proteins and visualizing them with ^{125}I-labeled avidin or avidin-based chromophores. Avidin-biotin complex (ABC) detection kits are available for isolated proteins (described in this protocol) as well as for proteins on fixed tissue, tissue sections, and microplate wells.

Materials

Biotinylated experimental protein (see Basic Protocols 1 to 4)
Biotinylated protein molecular weight standards (Bio-Rad or Vector Labs)
3% (w/v) BSA in PBS (*APPENDIX 2E*)
Nitrocellulose membrane
0.5 µCi ^{125}I-Bolton-Hunter-labeled avidin (DuPont NEN) in 5 ml avidin buffer (see recipe) *or* Vectastain ABC kit (Vector Labs) *or* ImmunoPure ABC kit (Pierce)

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*), electroblotting (*UNIT 10.7*), and detecting proteins in immunoblots (*UNIT 10.10*)

1. Electrophorese biotinylated experimental protein and protein molecular weight standards in separate lanes on a one-dimensional SDS–polyacrylamide gel (*UNIT 10.1*)

2. Electroblot proteins onto a nitrocellulose membrane (*UNIT 10.7*) and block ~1 hr at 22°C with 3% BSA in PBS.

   Note that the milk-based blocking solution described in UNIT 10.10 should not be used because milk contains free biotin that can lead to unacceptably high backgrounds.

3a. *To visualize biotinylated proteins by labeling with radioactive avidin*: Incubate the blot in 0.5 µCi of ^{125}I-Bolton-Hunter-labeled avidin (Mock, 1990) in 5 ml avidin buffer for 2 hr at 22°C. Wash the membrane extensively in the same buffer without avidin or BSA, air dry, and expose to film (Goodloe-Holland and Luna, 1987).

3b. *To visualize biotinylated proteins with avidin-linked chromophores in the Vectastain ABC kit or ImmunoPure ABC kit*: Follow the manufacturer's instructions.

   The ABC procedure is essentially identical to that described in UNIT 10.10, except that steps 5 and 6 (incubation with a biotinylated secondary antibody) are omitted.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Avidin buffer
150 mM NaCl
4% (w/v) BSA
0.02% (v/v) Triton X-100
50 mM sodium borate
Adjust to pH 10.0 to 10.5 with 1 M NaOH
Store ≤2 months at 4°C

SCU IASI CENTRAL UNIVERSITY LIBRARY

Coupling buffer

Select one of the following buffers (see Critical Parameters for discussion of practical considerations and APPENDIX 2E for recipes):

100 mM $NaHCO_3$, pH 8.5
100 mM Na_2CO_3, pH 8.5
100 mM sodium phosphate buffer, pH 7.5
100 mM sodium phosphate buffer, pH 8.0
100 mM sodium phosphate buffer/5 mM EDTA, pH 8.3

Physiological buffer, pH 7.4

7 g NaCl (120 mM final)
22 ml 0.2 M KH_2PO_4 (*APPENDIX 2E*; 4.4 mM final)
103 ml 0.2 M Na_2HPO_4 (*APPENDIX 2E*; 20.6 mM final)
Add H_2O to 1 liter
Check pH (it should be ~7.4; if not, check concentrations of stock solutions)
Store ≤2 months at 4°C or ≤2 weeks at room temperature

COMMENTARY

Background Information

The extraordinarily high affinity of avidin for biotin was first discovered as the explanation for a deficiency disease caused by ingestion of large numbers of raw eggs. In this disease, which is characterized by skin lesions and growth retardation, the essential vitamin biotin is depleted because it is sequestered by avidin in egg whites (György et al., 1941). The dissociation constant at each of the four biotin-binding sites of avidin is ~10^{-15} M (Green, 1975), making the avidin-biotin interaction one of the strongest known noncovalent biological interactions. Retention of high-affinity binding after derivatization of the biotin carboxyl group with a wide variety of agents has led to the development of numerous assays involving the recovery or detection of biotin-tagged moieties with avidin or other high-affinity biotin-binding proteins described below (reviewed by Wilchek and Bayer, 1990 and Savage et al., 1992).

Selection of biotinylation reagent

There are a number of reagents that can be used for biotinylation (see Table 3.6.1). Selection of reagent depends on the intended uses for the biotinylated product. For instance, most enzyme activities are independent of protein modification on carbohydrate residues, which generally can be assumed to be innocuous; labeling of cysteinyls, however, which generally can be assumed to be innocuous, may disrupt disulfide bonds that are important for structure or activity, so biotinylation of cysteinyls is not recommended for extracellular proteins. Biotinylation of lysyl groups works well with many proteins, including those on surfaces of intact cells, but it also may lead to loss of protein function.

N-Hydroxysuccinimide (NHS) esters of biotin react readily at alkaline pH with any accessible, unprotonated primary amine, such as ε-amino groups of lysyl side chains in proteins (see Basic Protocol 1 and Basic Protocol 2). Because lysine is abundant in proteins, accounting for ~7% of total amino acid residues in an average protein, NHS-biotin and its derivatives (see Table 3.6.1) are the most commonly used biotinylating agents. Modifications of NHS-biotin include addition of a sulfonate group to improve solubility in water (sulfo-NHS-biotin), insertion of a long-chain (LC) spacer arm between the biotin moiety and the reactive ester to increase the availability of the biotin at the protein surface (NHS-LC-biotin), and addition of a cleavable site, such as a disulfide bond (NHS-SS-biotin), to permit release of the protein from biotin using reducing agents such as 2-mercaptoethanol (2-ME) or dithiothreitol (DTT). A cleavable disulfide bond is particularly useful for eluting biotinylated proteins from an affinity matrix containing covalently bound avidin because the exceptionally high affinity (K_d ~10^{-15} M) and correspondingly low off-rate of the biotin-avidin interaction complicates elution with chaotropes (but see below).

Several commercially available reagents (see Table 3.6.1) permit biotinylation of proteins on cysteinyl residues. At physiological pH (pH 7.5 to 8.0), iodoacetyl-LC-biotin, biotin-LC-maleimide, maleimidopropionyl biocytin, biotin-HPDP, and biotin-maleimide all react

BOSTON CENTRAL UNIVERSITY LIBRARY

primarily with free sulfhydryl groups on reduced cysteines. For cytoplasmic proteins, in particular, labeling with cysteine-directed reagents may permit retention of functions lost after derivatization of lysines. However, these reagents may not be advisable for labeling extracellular or membrane proteins with conformations stabilized by disulfide bonds because reduction of at least some disulfides with thiol-reducing agents (e.g., 50 mM DTT or 100 mM 2-ME) is generally necessary for efficient labeling (see Support Protocol 1).

Biotin hydrazides (see Table 3.6.1), especially those with a six-carbon spacer (biotin-LC-hydrazide), can be excellent reagents for biotinylating galactose and *N*-acetylgalactosamine residues on extracellular and transmembrane proteins (see Basic Protocol 4). The biotin hydrazide reagents covalently bind to aldehydes that have been generated from vicinal carbohydrate hydroxyls (*cis*-diols) by mild chemical or enzyme oxidation. The resulting hydrazone bonds are stable between pH 2 and 10. Sodium *m*-periodate provides a fast, effective means for oxidizing *cis*-diols on many soluble proteins, including immunoglobulins. However, the reagent also can react with serines, threonines, and tyrosines. Consequently, periodate oxidation sometimes affects protein biological activity and often alters migration on SDS–polyacrylamide gels. If these side effects cannot be tolerated (e.g., when identifying cell-surface proteins by biotinylation), then galactose and *N*-acetylgalactosamine residues should be exposed and oxidized by treating the cells on membranes with a combination of neuraminidase and galactose oxidase.

Detection of biotinylated proteins

A large number of biotinylated proteins and biotin-detection reagents are commercially available, either as individual compounds or in the form of kits. For instance, biotinylated dipalmitoyl-L-α-phosphatidylethanolamine (Pierce) can be used as a phospholipid tracer in membrane reconstitution studies. Also, biotinylated insulin (Sigma), testosterone (Sigma), and lectins (Pierce or Vector Labs) can be used to identify and/or purify their cellular receptors. Avidin-biotin complex (ABC) detection kits (e.g., Vectastain ABC kit, Vector Labs, and ImmunoPure ABC kit, Pierce; see Support Protocol 2) amplify even weak signals by taking advantage of the multivalency of avidin for biotin and the incorporation of multiple biotins into most biotinylated antibodies and enzymes. Other, similar kits from Vector Labs and Pierce permit visualization of biotinylated proteins on fixed tissue, thin microscopic sections, and microtiter plate wells, allowing use of the biotin-avidin high-affinity interaction in immunofluorescence microscopy, electron microscopy, and enzymatic assays, as well as for protein size determinations.

The molecular masses of experimentally obtained biotinylated proteins are best estimated by comparison with the sizes of biotinylated protein molecular weight markers (Bio-Rad or Vector Labs). Because the addition of one or a few of the small biotin molecules does not significantly increase the mass of most proteins, the sizes of biotinylated protein standards are considered identical to those of prestained or unstained markers. Additionally, the markers are visualized with the same detection methods as the proteins of interest, permitting direct size comparisons.

Nonspecific, biotin-independent binding to avidin can usually be controlled by employing buffers of pH 10.0 to 10.5, which is near the avidin isoelectric point (pI 10.0). Use of such buffers minimizes electrostatic interactions but causes little or no diminution of the avidin-biotin affinity. When high-pH buffers are undesirable, avidin should be replaced either by a neutralized, deglycosylated form of avidin (Neutralite Avidin from Molecular Probes or Southern Biotechnology, or NeutrAvidin from Pierce) or by streptavidin (Pierce, Sigma, and Vector Labs). Streptavidin is a tetrameric protein from the culture broth of *Streptomyces avidinii*; it binds four biotin molecules per mole with an affinity similar to that of avidin. Both neutralized avidin (pI 6.3) and streptavidin (pI 5.0) exhibit fewer nonspecific electrostatic interactions at physiological pH and ionic strength. Streptavidin does, however, contain a sequence that causes biotin-independent binding to integrins and related cell-surface fibronectin receptors under certain conditions (Alon et al., 1990). Many commercially available derivatives of streptavidin (Pierce, Sigma, and Vector Labs) and neutralized avidin (Pierce and Southern Biotechnology) may be used in microscopic and enzymatic assays as a replacement for underivatized avidin.

Recovery of biotinylated proteins

Paradoxically, the high affinity of the biotin-avidin interaction can be a disadvantage during affinity purification of biotinylated proteins. Biotinylated proteins bind tightly to avidin affinity matrices (Molecular Probes, Pierce, Sigma, or Vector Labs) under almost

PCU IASTATE UNIVERSITY LIBRARY

any physiological conditions, but the high-affinity binding is reversed only by harsh denaturants, such as 8 M guanidine·HCl, pH 1.5, or by boiling with 2% (w/v) SDS. Further, the large positive charge on avidin (pI 10.0) and its carbohydrate residues promote nonspecific binding of many negatively charged proteins and some lectins to avidin affinity columns. For this reason, matrices containing bound NeutrAvidin (Pierce) or streptavidin (Molecular Probes, Pierce, Sigma, and Vector Labs) are often used for the affinity purification of biotinylated proteins, despite the significantly greater expense of these matrices.

For preparative purification of biotinylated proteins by affinity chromatography, whether it be on columns containing bound avidin, streptavidin, or neutralized avidin, there are two approaches. The first approach is to accept that the biotin-avidin interaction is essentially irreversible and to attach the biotin label to a dispensable moiety that is left on the column after eluting the protein of interest. For instance, biotinylated monoclonal antibodies can be used to purify individual membrane antigens (Updyke and Nicolson, 1984), and biotinylated insulin and other ligands have been used to purify their specific receptors by avidin or streptavidin affinity chromatography (Kohanski and Lane, 1985/1986; Kraehenbuhl and Bonnard, 1990). Similarly, disulfide bond reduction (see Support Protocol 1) will release proteins that have been biotinylated with NHS-SS-biotin (Chia and Luna, 1989) or with biotin-HPDP (see Table 3.6.1).

The second approach to purifying native proteins is to moderate the avidin-biotin binding affinity by modifying either the bound avidin or the biotin. For example, "monomeric" avidin columns (Pierce or Sigma), generated by treating immobilized avidin with 0.1 M guanidine·HCl, pH 1.5, contain one lower-affinity biotin-binding site (K_d ~5×10^{-8} M) per bound avidin molecule. After presaturating residual high-affinity sites on the column with excess free biotin, biotin-containing proteins can be bound to the columns and then eluted with either 0.1 M glycine, pH 2.0 to 2.8, or 2 mM free biotin (Henrikson et al., 1979; Kohanski and Lane, 1990).

A converse strategy is to derivatize protein lysines with 2-iminobiotin-NHS (see Table 3.6.1), a guanido analog of NHS-biotin that exhibits a strong pH dependence in interactions with avidin and streptavidin (Heney and Orr, 1981). At pH 9.5, iminobiotin binds avidin/streptavidin with characteristically high affinity. However, proteins labeled with iminobiotin can be eluted from avidin columns with either 0.1 M acetic acid or 50 mM ammonium acetate/0.5 M NaCl, pH 4.0. Similarly, columns containing bound 2-iminobiotin (Molecular Probes, Pierce, or Sigma) can be used to purify avidin, streptavidin, and their derivatives. For further information on affinity chromatography, see Chapter 9.

Critical Parameters and Troubleshooting

The most critical aspects of protein biotinylation are the storage and solubilization conditions for the biotinylating reagent. This is particularly critical with NHS derivatives, which contain succinimide ester groups that are very reactive with water. The NHS-biotins should be stored in the presence of desiccant at –20°C, as recommended by the suppliers, and the reagent bottles should be warmed completely to room temperature before opening to minimize condensation. Similarly, it is critical to mix the compounds with the protein to be labeled as soon as possible after dissolving the reagents in aqueous solutions.

Most other parameters are less critical. Although the rate of protein biotinylation is influenced by temperature, short incubations (e.g., 30 min to 1 hr) at either 0° to 4°C or room temperature are usually satisfactory. The extent of biotinylation is determined by the amounts of reagents mixed and the reaction time. Biotin-to-protein molar ratios from 1:1 to 60:1 and reaction times up to 18 hr have been described. Because biotinylation is directed at selected groups in proteins (e.g., labeling ranges from a single cysteinyl or carbohydrate to some or all lysyls, which make up 7 of 100 residues in a typical protein), the choice of biotinylating reagent is often more critical than the reaction conditions. If an enzyme loses activity during biotinylation and if a change of biotinylating reagent is contraindicated, one strategy is to protect the enzyme's active site by adding excess substrate during the biotinylation reaction.

The medium in which biotinylation takes place must be free of components that will interfere with the reaction chemistry. Coupling buffers for NHS-biotin reagents (see Basic Protocols 1 and 2) must be free of ingredients with primary amines. Excess thiol-reducing agent (see Support Protocol 1) must be removed for efficient use of sulfhydryl-directed reagents (see Basic Protocol 3), and excess *cis*-diol-oxidizing agent must be eliminated for labeling of protein carbohydrates (see Basic Proto-

PUBLISHED BY CENTRAL UNIVERSITY LIBRARY

col 4). Finally, pH is an important factor. Lysyl-directed reagents require unprotonated amines (e.g., pH 7.5 to 8.5), whereas maleimide-based thiol-directed reagents undergo side reactions with amines at pH 7.5 or above and thus need more acid conditions (e.g., pH 6.5 to 7.5) if specificity for cysteines is required. Carbohydrate-directed reagents also require acid conditions (e.g., pH 5.5) for specificity.

Reagent solubility may also be a factor. If the protein of interest is sensitive to organic solvents, such as DMF and DMSO, use the water-soluble sulfo-NHS-biotin for labeling lysyl groups. If highest affinity binding is not needed, biotinylating reagents lacking the LC spacer may be more water soluble.

When biotinylating a subset of proteins in a complex mixture—for example, when determining which plasma membrane proteins are accessible to a cell-surface biotinylating reagent—it is useful to include two controls in each experiment. First, some cells should always be carried through the labeling and fractionation procedure in the absence of biotinylating reagent, in order to identify endogenously biotinylated proteins, which are involved in many intracellular carboxylation reactions. Second, to help identify proteins that may be exposed by cell breakage during manipulation, cytoplasmically oriented proteins should be labeled with biotinylating reagent either after deliberate mechanical lysis in the presence of 0.2% (v/v) Triton X-100 or after plasma membrane purification. The group of biotinylated cytoplasmically oriented proteins generated with this control should differ dramatically from that observed after biotinylation of unbroken cells. If the two SDS-PAGE profiles appear similar, then the biotinylating procedure should be repeated under conditions that are less conducive to cell lysis (e.g., biotinylate cells that are still attached to tissue culture plastic).

Because proteins differ dramatically in the number of surface-accessible reactive groups and how they are affected by labeling with a foreign tag such as biotin, it is hard to predict which biotinylating compound will work best for a particular protein. However, biotinylation chemistries are similar to those used for a variety of other compounds, including fluorescent labels. Thus, if a literature reference exists for fluorescently labeling a protein of choice, the most expeditious strategy is to obtain a biotinylating compound with the same reactive group (i.e., succinimide ester, maleimide, or hydrazide moiety) and to follow the published reaction conditions for labeling the protein.

If no information is available about accessible groups or about the functional consequences of a particular derivatization strategy, it is a good idea to try different types of biotinylating reagents at different reagent-to-protein ratios. Reagent/protein molar ratios of 1:1 and 20:1 are usually good starting points. Preliminary characterization of protein functionality should suggest which of the reaction chemistries is superior and whether a higher or lower stoichiometry yields the most efficient labeling.

Anticipated Results

The biotinylation protocols in this unit should biotinylate most proteins. The extent of reaction and the degree to which the protein retains its endogenous structure and activity will vary but may be predictable if other investigators have used a similar chemistry to derivatize the protein with a "tag" other than biotin.

Soluble proteins biotinylated by these procedures have been used as ligands in binding assays in vitro and as tracers for cellular processes including endocytosis, membrane trafficking, and cytoskeletal remodeling (reviewed in Savage et al., 1992, and Wilchek and Bayer, 1990). Cell surface–biotinylated membrane proteins can be used similarly as markers for plasma membrane proteins in biochemical fractionations and in membrane recycling assays. In addition, biotinylated protein(s) can be purified, either alone or in combination with receptors or other binding partners, on avidin or streptavidin affinity columns.

Time Considerations

Labeling reactions involving succinimide esters (NHS-biotin and analogs) and those involving biotin-hydrazides are relatively rapid and usually complete within 1 to 2 hr at 4°C. Reactions of sulfhydryls with maleimides and iodoacetamides are slower, particularly when carried out in lower-pH buffers that optimize reaction selectivity of protein sulfhydryls over amino groups. Although these reactions can be terminated after only 2 to 4 hr, reaction extents are usually increased by longer incubations (e.g., overnight at 4°C).

Similarly, depending on the choice of separation technique, removal of unincorporated biotinylating reagent from labeled protein can be either rapid or slow, with gel-filtration chromatography and centrifugal filtration requiring

UCSF LIBRARY

no more than 1 to 2 hr, whereas dialysis may take 1 to 3 days. Separation steps also are often dictated more by the properties of the protein being labeled than by investigator preference. Some proteins become irreversibly aggregated if subjected to centrifugal filtration; other proteins are denatured or proteolyzed during prolonged dialysis. Here again, it helps to know what other investigators have done successfully with the particular protein under investigation.

Literature Cited

Alon, R., Bayer, E.A., and Wilchek, M. 1990. Streptavidin contains an RYD sequence which mimics the RGD receptor domain of fibronectin. *Biochem. Biophys. Res. Commun.* 170:1236-1241.

Bayer, E.A., Ben-Hur, H., and Wilchek, M. 1988. Biocytin hydrazide—A selective label for sialic acids, galactose, and other sugars in glycoconjugates using avidin-biotin technology. *Anal. Biochem.* 170:271-281.

Chia, C.P. and Luna, E.J. 1989. Phagocytosis in *Dictyostelium discoideum* is inhibited by antibodies directed primarily against common carbohydrate epitopes of a major cell-surface plasma membrane glycoprotein. *Exp. Cell Res.* 181:11-26.

Goodloe-Holland, C.M. and Luna, E.J. 1987. Purification and characterization of *Dictyostelium discoideum* plasma membranes. *Methods Cell Biol.* 28:103-128.

Green, N.M. 1975. Avidin. *Adv. Protein Chem.* 29:85-133.

Gyöargy, P., Rose, C.S., Eakin, R.E., Snell, E.E., and Williams, R.J. 1941. Egg-white injury as the result of nonabsorption or inactivation of biotin. *Science* 93:477-478.

Heney, G. and Orr, G.A. 1981. The purification of avidin and its derivatives on 2-iminobiotin-6-aminohexyl-Sepharose 4B. *Anal. Biochem.* 114:92-96.

Henrikson, K.P., Allen, S.H.G., and Maloy, W.L. 1979. An avidin monomer affinity column for the purification of biotin-containing enzymes. *Anal. Biochem.* 94:366-370.

Ingalls, H.M., Goodloe-Holland, C.M., and Luna, E.J. 1986. Junctional plasma membrane domains isolated from aggregating *Dictyostelium discoideum amebae. Proc. Natl. Acad. Sci. U.S.A.* 83:4779-4783.

Kohanski, R.A. and Lane, M.D. 1985/1986. Receptor affinity chromatography. *Ann. N.Y. Acad. Sci.* 447:373-385.

Kohanski, R.A. and Lane, M.D. 1990. Monovalent avidin affinity columns. *Methods Enzymol.* 184:194-200.

Kraehenbuhl, J.-P. and Bonnard, C. 1990. Purification and characterization of membrane proteins. *Methods Enzymol.* 184:629-641.

Mock, D.M. 1990. Sequential solid-phase assay for biotin based on ^{125}I-labeled avidin. *Methods Enzymol.* 184:224-233.

O'Shannessy, D.J., Doberson, M.J., and Quarles, R.H. 1984. A novel procedure for labeling immunoglobulins by conjugation to oligosaccharide moieties. *Immunol. Lett.* 8:273-277.

Savage, M.D., Mattson, G., Desai, S., Nielander, G.W., Morgensen, S., and Conklin, E.J. 1992. Avidin-Biotin Chemistry: A Handbook. Pierce Chemical Co., Rockford, Ill.

Updyke, T.V. and Nicolson, G.L. 1984. Immunoaffinity isolation of membrane antigens with biotinylated monoclonal antibodies and immobilized streptavidin matrices. *J. Immunol. Methods* 73:83-95.

Wilchek, M. and Bayer, E.A. (eds.) 1990. Avidin-biotin technology. *Methods Enzymol.* 184:1-746.

Contributed by Elizabeth J. Luna
Worcester Foundation for Biomedical Research
Shrewsbury, Massachusetts

BCU IASI / CENTRAL UNIVERSITY LIBRARY

EXTRACTION, STABILIZATION, AND CONCENTRATION

4

INTRODUCTION 4.0.1

4.1 Overview of Cell Fractionation 4.1.1
Basic Principles of Centrifugation 4.1.2
Instrumentation 4.1.2
Fractionation Media 4.1.5
Evaluation of Fractionation 4.1.7
Definitive Procedures 4.1.8

4.2 Purification of Organelles from Mammalian Cells 4.2.1
Strategic Planning 4.2.1
Basic Protocol 1: Differentiation Centrifugation by Velocity 4.2.7
Basic Protocol 2: Differentiation Centrifugation by Equilibrium 4.2.9
Alternate Protocol 1: Rate Zonal Centrifugation Using Sucrose 4.2.12
Alternate Protocol 2: Gradient Fractionation Using Percoll 4.2.14
Alternate Protocol 3: Gradient Fractionation Using D_2O 4.2.18
Alternate Protocol 4: Gradient Fractionation Using Nonionic Iodinated Solutes 4.2.20
Basic Protocol 3: Gel Filtration to Isolate Secretory Vesicles 4.2.22
Basic Protocol 4: Preparative Electrophoretic Separation 4.2.24
Alternate Protocol 5: Electrophoretic Separation on a Density Gradient 4.2.27
Basic Protocol 5: Solid-Phase Immunoadsorption 4.2.30
Basic Protocol 6: Purification by Lectin Adsorption 4.2.33
Basic Protocol 7: Density Shift Using Digitonin 4.2.35
Basic Protocol 8: Density Shift Using Colloidal Gold Conjugates 4.2.38
Basic Protocol 9: Extraction of Extrinsic Proteins from Membranes Using Sodium Carbonate 4.2.42
Reagents and Solutions 4.2.43
Commentary 4.2.45

4.3 Subcellular Fractionation of Tissue Culture Cells 4.3.1
Basic Protocol 1: Internalization of Fluid-Phase Marker 4.3.1
Basic Protocol 2: Implantation and Internalization of VSV G Protein 4.3.3
Support Protocol 1: Preparation of VSV Stocks 4.3.5
Basic Protocol 3: Homogenization of Tissue Culture Cells 4.3.6
Support Protocol 2: Determination of Horseradish Peroxidase Activity 4.3.7
Basic Protocol 4: Flotation-Gradient Fractionation of Tissue Culture Homogenates 4.3.8
Basic Protocol 5: Immunoisolation of Endosomal Fractions 4.3.9
Support Protocol 3: Preparation of Tissue Culture Cells for Subcellular Fractionation 4.3.10
Reagents and Solutions 4.3.10
Commentary 4.3.12

continued

BCU IAS CENTRAL UNIVERSITY LIBRARY

Copyright © 1997 by John Wiley & Sons, Inc.

4.4 Desalting, Concentration, and Buffer Exchange by Dialysis and Ultrafiltration 4.4.1
Basic Protocol 1: Desalting and Buffer Exchange by Dialysis Through Regenerated Cellulose Tubing 4.4.1
Basic Protocol 2: Concentration or Dialysis by Ultrafiltration Through Asymmetric Membrane Disks 4.4.4
Alternate Protocol 1: Diafiltration or Concentration by Tangential-Flow Ultrafiltration 4.4.6
Alternate Protocol 2: Microscale Concentration and Desalting with Centrifugal Ultrafilters 4.4.8
Commentary 4.4.9

4.5 Selective Precipitation of Proteins 4.5.1
Strategic Planning 4.5.2
Basic Protocol 1: Selective Precipitation by Salting Out 4.5.5
Alternate Protocol 1: Selective Precipitation by Stepwise Salting Out 4.5.8
Basic Protocol 2: Selective Precipitation by Isoionic Precipitation: Column Method 4.5.9
Alternate Protocol 2: Selective Precipitation by Isoionic Precipitation: Dialysis Method 4.5.12
Basic Protocol 3: Selective Precipitation Using a Two-Carbon (C_2) Organic Cosolvent 4.5.13
Basic Protocol 4: Selective Precipitation Using C_4 and C_5 Organic Cosolvents 4.5.14
Basic Protocol 5: Selective Precipitation Using Protein Exclusion and Crowding Agents and Osmolytes 4.5.15
Basic Protocol 6: Selective Precipitation Using Synthetic and Semisynthetic Polyelectrolytes 4.5.16
Basic Protocol 7: Selective Precipitation Using Metallic and Polyphenolic Heteropolyanions 4.5.20
Basic Protocol 8: Selective Precipitation Using Hydrophobic Ion Pairing (HIP) Entanglement Ligands 4.5.20
Basic Protocol 9: Selective Precipitation by Matrix-Stacking Ligand Coprecipitation 4.5.21
Basic Protocol 10: Selective Precipitation Using Di- and Trivalent Metal Cation Precipitants 4.5.25
Commentary 4.5.27

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 4
Extraction, Stabilization, and Concentration

INTRODUCTION

Sources of proteins to be used for biological assays and for purification purposes are diverse. Recombinant DNA technology has allowed the production of many proteins in forms suitable for biological assays as well as structural analysis; methods for purifying and characterizing proteins from recombinant sources are described in Chapters 5 to 7. In many cases, gene cloning methods lead to the identification and production of proteins even when they have not previously been purified from conventional sources. This type of approach often yields large amounts of protein, which may then be used to generate immunological reagents that can be used in localization studies or to explore the protein's expression patterns in normal tissues or tissue fractions. It remains true that important information about the biological and functional properties of proteins may be inferred from their intracellular distribution and subcellular localization.

For other polypeptides that have not yet been obtained in recombinant form, subcellular fractions may still constitute a convenient, if not the only, source of a protein of interest. Because so many proteins are expressed in a tissue- or cell-type-specific fashion, obtaining the appropriate starting material is not always a trivial matter. Cultured cell lines may, but do not necessarily, express exactly the same protein profile as the tissue that gave rise to the cell lines that happen to be available, and for a number of differentiated cell types no culture methods have yet been described. For this reason, access to subcellular fractions prepared from tissue can be critical. Moreover, the ability to demonstrate the presence or activity of polypeptides of interest in defined subcellular locations and to retrieve them for further study is a key element of protein characterization. Thus methods for fractionating cells, though more commonly considered part of cell biology rather than molecular biology or protein chemistry, remain critical tools in the arsenal of any scientist attempting to isolate and characterize a protein.

The first part of this chapter presents methods for subcellular fractionation of tissues and cultured cells. *UNIT 4.1* discusses a number of basic considerations that should guide the development of specific subcellular fractionation procedures. *UNIT 4.2* provides a number of protocols for preparing the most commonly used subcellular fractions from readily available tissues. Also covered are criteria for assessing the purity of the resulting subcellular fractions and methods for further treatment of membrane fractions to discriminate between integral and peripheral membrane proteins. The details of these protocols apply specifically to the examples described, but are also intended as general guides. The procedures can be modified and optimized for particular applications and for use with a wide range of tissue and cell types. *UNIT 4.3* applies immunochemical methods to obtaining subcellular fractions. These protocols, too, should serve as guides and may be adapted for application to the specific problems at hand.

Most often, the process of protein extraction and purification results in high concentration of low-molecular-weight contaminants. Desalting, buffer exchange, and protein concentration by dialysis and ultrafiltration are discussed in *UNIT 4.4*.

UNIT 4.5 provides a series of tools for the protein chemist that are familiar to many, but have not recently been compiled in an accessible manner: protocols for the selective

Contributed by Hidde L. Ploegh
Current Protocols in Protein Science (1997) 4.0.1-4.0.2
Copyright © 1997 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

precipitation of proteins. In many cases, proteins of interest must be recovered from dilute solutions containing protein and nonprotein contaminants. Practical problems for downstream processing may arise because of the volumes to be handled. *UNIT 4.5* offers some practical solutions to these problems.

Hidde L. Ploegh

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of Cell Fractionation

Cell fractionation has enjoyed widespread use among cell biologists for half a century. It continues to be a fruitful, if not essential, approach in the reductionistic efforts to define the composition and functions of the multiple compartments in eukaryotic cells. It also provides the essential ingredients for the increasing number of cell-free assays now being used in test-tube reconstructions of complex cellular events involving intercompartmental interactions. Much of the knowledge regarding the composition and function of cell organelles has resulted from fractionation of mammalian tissues, where cells are both abundant and highly differentiated (and thus organelle-rich). However, with new techniques in molecular biology and widened interest in combining studies of intact cells and functional reconstitution, interest in fractionation has spread to cultured cell lines and genetically tractable lower eukaryotic cells.

The goals of cell fractionation often differ depending on the nature of the experiments being conducted. In preparative procedures, in which the intent is to isolate quantities of a particular cell organelle for further study or for subfractionation, the emphasis is on purity and (secondarily) on yield. In analytical experiments, in which the intent is not isolation of organelles but evaluation of associations of selected macromolecules with particular organelles, the emphasis is on using one-step procedures that result in different distributions of various organelles (as defined by marker activities) rather than on separating organelles outright. Finally, in preparing organelles for cell-free reconstitution, the goal is to maintain them in a functional state. The investigator generally has developed a specific assay for intercompartmental interaction that does not rely on organelle purity, so the extent of contamination by irrelevant organelles is less important.

The separation of distinct organelles during cell fractionation results from their differing physical properties—size (and shape), buoyant density, and surface charge density—which reflect their differing compositions. Particular fractionation techniques capitalize on one or more of these properties. For example, gel filtration separates on the basis of size, centrifugation separates on the basis of size and density, and electrophoresis separates on the basis of surface charge density. As knowledge of the specific composition of particular organelles has developed, it has become possible to apply newer techniques such as affinity chromatography and selective density-shift perturbation. In *UNIT 4.2*, examples are given of fractionation procedures based on each of these approaches. Some are preparative and others are analytical. Although an effort has been made to present procedures that are widely applicable and accessible and that do not require highly specialized and expensive equipment, the procedures may have to be modified to adapt them to individual needs. Even isolation of a particular kind of organelle from different tissue or cell sources may necessitate adjustment of fractionation conditions. This also means that in addition to isolating the organelle of interest, you should also plan on confirming the identity of what you have isolated.

Centrifugation is the most widely used procedure in cell fractionation, as indicated in *UNIT 4.2*. It is the only approach commonly used to separate crude tissue homogenates (often having quite large volumes) into subfractions as starting material for more refined purification procedures. Further, the technology available, using rotors with a variety of geometries and diverse media that enable separation according to size, density, or both, now routinely permits refined separations on volumes ranging from submilliliter to several liters. No other technology is this versatile. Gel filtration is limited by the pore sizes of available resins, such that only regularly shaped vesicles with diameters <100 to 200 nm can be purified away from larger or irregularly shaped organelles. Use of electrophoresis for organelle purification, especially at the preparative level, is relatively recent, and, as with gel filtration, its success relies on generating starting material by centrifugation. Although it shows promise for purifying selected organelles (e.g., endosomes) that have been difficult to obtain otherwise, surface charge densities on most organelles may not be different enough (or able to be manipulated sufficiently) to make this a versatile procedure. For these reasons, the remaining comments in this overview focus primarily on fractionation of organelles by centrifugation. For more detailed coverage of each of the topics to be discussed, the reader is referred to the excellent book edited by Rickwood (1984).

Copyright © 1995 by John Wiley & Sons, Inc.

BOULDER / PENTRAL UNIVERSITY LIBRARY

BASIC PRINCIPLES OF CENTRIFUGATION

One of the best ways to understand the rationale for various cell fractionation procedures is to examine the Svedberg equation (Equation 4.1.1), which describes mathematically the sedimentation of a spherical particle in a viscous fluid:

$$\frac{dx/dt}{\omega^2 x} = \frac{2r^2(\rho_p - \rho_m)}{9\eta}$$

Equation 4.1.1

where x is the distance from axis of rotation, t is time, ω is the angular velocity, r is the radius of the particle, ρ_p and ρ_m are the density of the particle and medium, and η is the viscosity of the medium.

This equation states that the sedimentation velocity, dx/dt, per unit centrifugal field $\omega^2 x$ (which is set by the instrumentation) increases with the square of the particle radius and the difference in density between the particle and the medium and decreases with increasing viscosity. In practical terms, when sedimentation is performed by differential velocity in a medium such as 0.25 M sucrose, which is less dense than all particles and has a low viscosity, the radius of the particle is the dominant factor that determines the sedimentation rate. In contrast, when isopycnic density gradient centrifugation is used, the density of the medium changes from top to bottom in the centrifugation tube, and the range of densities typically includes the buoyant densities of most particles (vesicles). Thus individual types of vesicles will sediment until $\rho_p = \rho_m$, and the vesicles will then stop at their isopycnic densities. Alternatively, in flotation procedures $\rho_p < \rho_m$ initially and dx/dt is negative, meaning that particles will move toward the axis of rotation (i.e., toward the top of the tube). Evidently, the equation also indicates that sedimentation or flotation, as the case may be, is slowed in media of higher viscosity. As discussed below, various media used in centrifugation are distinguished by their viscosities (and their osmolarities) at high densities.

The Svedberg equation can also be used to discuss three other basic relationships encountered in using centrifugation. First, the velocity of a particle per unit centrifugal field (left side of the equation) is the sedimentation coefficient. Its dimensions are in seconds, and it is generally reported in units known as Svedbergs (S). 1 S = 10^{-13} sec, the same order of magnitude as the sedimentation coefficient of several biological macromolecules (e.g., 5S RNA). Second, the centrifugal field, $\omega^2 x$ is usually normalized to gravity and is expressed in the literature as a relative centrifugal force (RCF = $\omega^2 x/g$). Centrifugation conditions are best reported as $\times g$. RCF is usually expressed at the midpoint of the tube (g_{av}). The centrifugal field term includes the angular velocity in units of radians per second, which is easily converted to revolutions per minute (rpm), the unit most often given in the literature, as follows: RCF = 11.18 $r(N/1000)^2$, where N is rpm. Third, integration of the Svedberg equation over time yields the following relationship (Equation 4.1.2):

$$t = \frac{1}{s\,\omega^2} \ln\left(\frac{X_{max}}{X_{min}}\right) = \frac{k}{s}$$

$$\text{where } k = \frac{1}{\omega^2} \ln\left(\frac{X_{max}}{X_{min}}\right)$$

Equation 4.1.2

This is a measure of the clearance time, the time for a particle of given sedimentation coefficient to travel from the top to the bottom of the tube. Included in k are the geometry of the rotor (X_{max} and X_{min}) and the angular velocity. k-factors are reported in the printed information that accompanies most rotors. They are useful for estimating the length of centrifugation required to pellet a particular particle and for translating centrifugation conditions between two rotors. For a given particle or vesicle, k/t (rotor 1) = k/t(rotor 2).

INSTRUMENTATION

Several pieces of equipment must be available to carry out cell fractionation by centrifugation. These include homogenizers, centrifuges (generally one low-speed and one ultracentrifuge) with compatible rotors, a refractometer for measuring the refractive index of centrifugation media (before centrifugation) and gradient fractions (after centrifugation), a gradient-forming device for generating preformed density gradients, and a gradient-collecting device for unloading the gradients after the run. The refractometer, gradient-forming device, and gradient-collection device are very useful, if not essential, for density gradient and rate zonal sedimentation (*UNIT 4.2*).

Centrifuges. Most organelle purification procedures require that samples be maintained at ≤4°C, so the centrifuges used need to be

BCU IASCE LIBRARY UNIVERSITY

refrigerated. The low-speed centrifuge is used in the early steps of the typical purification procedure for sedimenting large particulates such as nuclei and unbroken cells, frequently from large volumes (e.g., *UNIT 4.2*). The ultracentrifuge is used for subsequent spins at higher speeds and often with smaller volumes. These instruments have a vacuum system as the rotors are spun in an evacuated chamber to reduce friction (and thus heating).

Rotors. Commonly used rotors come in three configurations: fixed-angle, vertical (or near-vertical), and swinging-bucket (Fig. 4.1.1). Continuous-flow and zonal rotors,

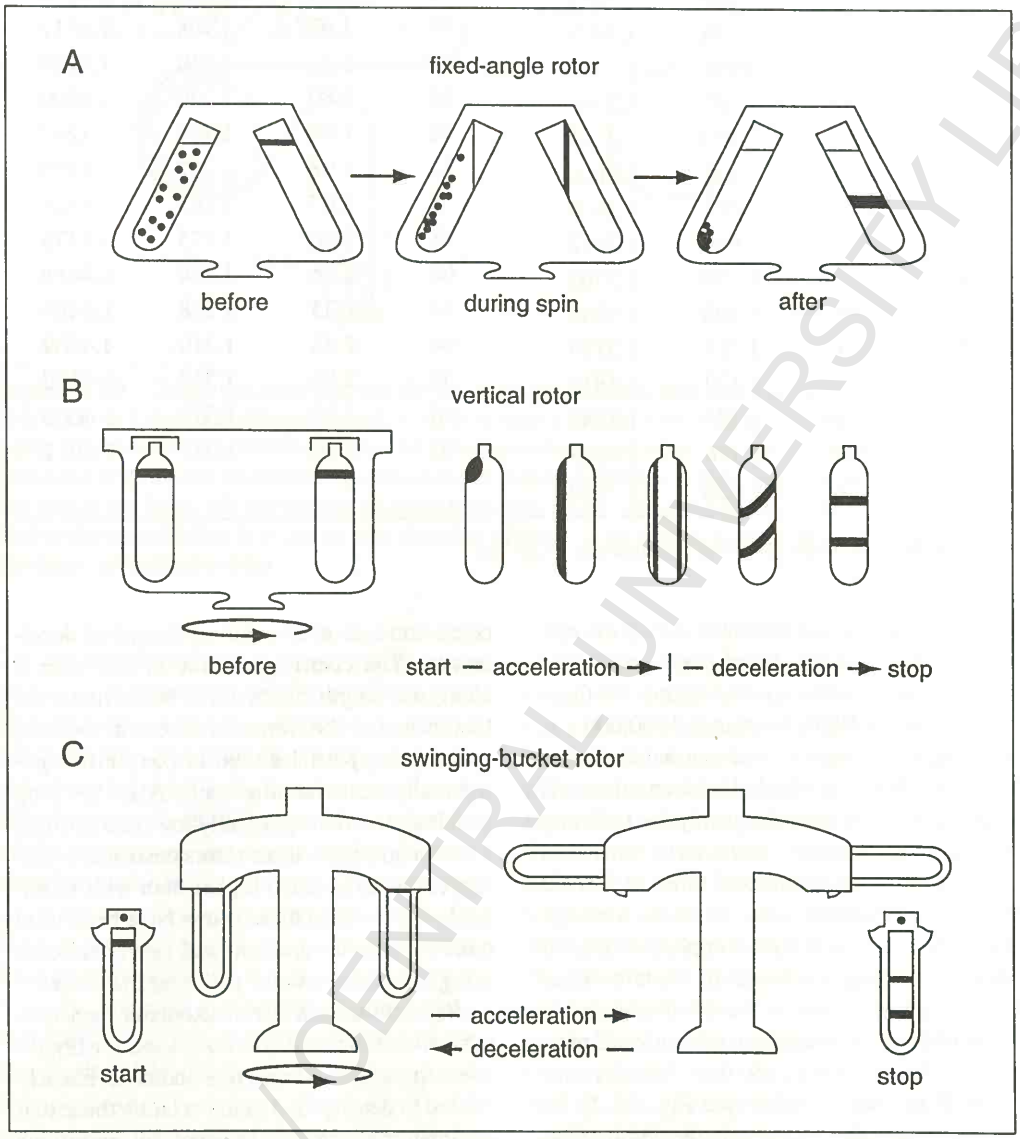


Figure 4.1.1 Rotors used in ultracentrifuges. (**A**) Fixed-angle rotor. In each profile, the tube on the left shows pelleting from a uniform suspension and the tube on the right shows sedimentation of a band layered at the top above a higher-density fluid. Sedimentation is radial (particles move toward the outer wall of the tube and then down this wall to the pellet). Note how the band in the right tube reorients during the spin and then again at the end of the run. (**B**) Vertical rotor. The rotor profile shows tubes with sample layered above a higher-density solution before a run. Note that during the run (successive tube profiles shown at right), the layer reorients along the inner wall of the tube and then the bands resolve vertically. During deceleration the bands reorient to a horizontal position. (**C**) Swinging-bucket rotor. Note the tube on the left in the upper profile has a sample layered at the top. The buckets are mounted vertically on the rotor. During the run, the buckets reorient to the horizontal position, and the bands separate along the length of the tube. Because sedimentation is radial, the particulates in the bands are more concentrated at the walls facing and away from the viewer than in the center of the tube. Toward the end of deceleration, the buckets reorient to the vertical position.

BOSTON UNIVERSITY / CENTRAL UNIVERSITY LIBRARY

Table 4.1.1 Density and Refractive Index of Aqueous Sucrose Solutions at 20°C[a]

Percent (w/w)	Molarity	Density	Refractive index	Percent (w/w)	Molarity	Density	Refractive index
2	0.06	1.006	1.3359	38	1.30	1.166	1.3958
4	0.12	1.014	1.3388	40	1.38	1.176	1.3997
6	0.18	1.022	1.3418	42	1.46	1.187	1.4036
8	0.24	1.030	1.3448	44	1.54	1.197	1.4076
10	0.30	1.038	1.3479	46	1.62	1.208	1.4117
12	0.37	1.046	1.3510	48	1.71	1.219	1.4158
14	0.43	1.055	1.3541	50	1.80	1.230	1.4200
16	0.50	1.064	1.3573	52	1.88	1.241	1.4242
18	0.56	1.072	1.3606	54	1.98	1.252	1.4285
20	0.63	1.081	1.3639	56	2.07	1.263	1.4329
22	0.70	1.090	1.3672	58	2.16	1.275	1.4373
24	0.77	1.099	1.3706	60	2.26	1.286	1.4418
26	0.84	1.108	1.3740	62	2.35	1.298	1.4464
28	0.91	1.118	1.3775	64	2.45	1.310	1.4509
30	0.99	1.127	1.3811	66	2.55	1.322	1.4558
32	1.06	1.137	1.3847	68	2.65	1.335	1.4605
34	1.14	1.146	1.3883	70	2.76	1.347	1.4651
36	1.22	1.156	1.3920				

[a]Modified from Hofmann (1977) by permission of ISCO, Inc.

which are loaded and unloaded during the run, are less widely used. Fixed-angle rotors emphasize speed and capacity. Many of these rotors generate RCFs in excess of 600,000 × *g*, meaning low *k*-factors, and can handle a few hundred milliliters of solution at one time. As a result, they are used frequently for pelleting fractions, particularly from large volumes. Most fixed-angle rotors hold tubes in a fixed incline substantially away from the vertical. When the centrifugal force is applied, it is radial rather than along the length of the tube. Thus the path length is across the tilted tube and is reasonably short; particles travel outward until they hit the outer wall and then "slide" down the wall to form a pellet (see Fig. 4.1.1). In vertical rotors, the path is only the width of the tube. As is evident from the figure, in both fixed-angle and vertical rotors the solution in the tubes reorients from the horizontal to the vertical during acceleration and reverses during deceleration. These rotors can be advantageous for density gradients (*UNIT 4.2*) because equilibrium is reached quite rapidly with the short path, and the separation of bands of different density increases as the contents of the tube reorient during deceleration. With swinging-bucket rotors, the tube and its contents reorient from vertical to horizontal during acceleration and reverse during the end of deceleration. The centrifugal force in this case is along the length of the tube. Note, however, that because the force is directed radially, sedimenting particles tend to concentrate peripherally near the tube walls. Also, the long path length and the generally lower centrifugal forces required for these rotors mean that *k*-factors are usually much higher than with fixed-angle rotors. Thus these rotors have been used more for density gradient and rate zonal centrifugations and less for pelleting fractions.

Refractometers. A refractometer measures refractive indices of solutions used for organelle separations. Refractive index is linearly related to density, and tables relating these two parameters over a large range of concentrations have been prepared for sucrose (Table 4.1.1) and are readily constructed by weighing aliquots of other accurately prepared gradient solutions. The measurement of refractive index/density is particularly useful for characterizing the position of organelles in gradients after centrifugation and for ensuring reproducibility of gradients from one experiment to another.

Gradient-forming devices. These devices consist of two chambers with an interconnecting channel along the bottom and an outlet from

BCU IASH CENTRAL UNIVERSITY LIBRARY

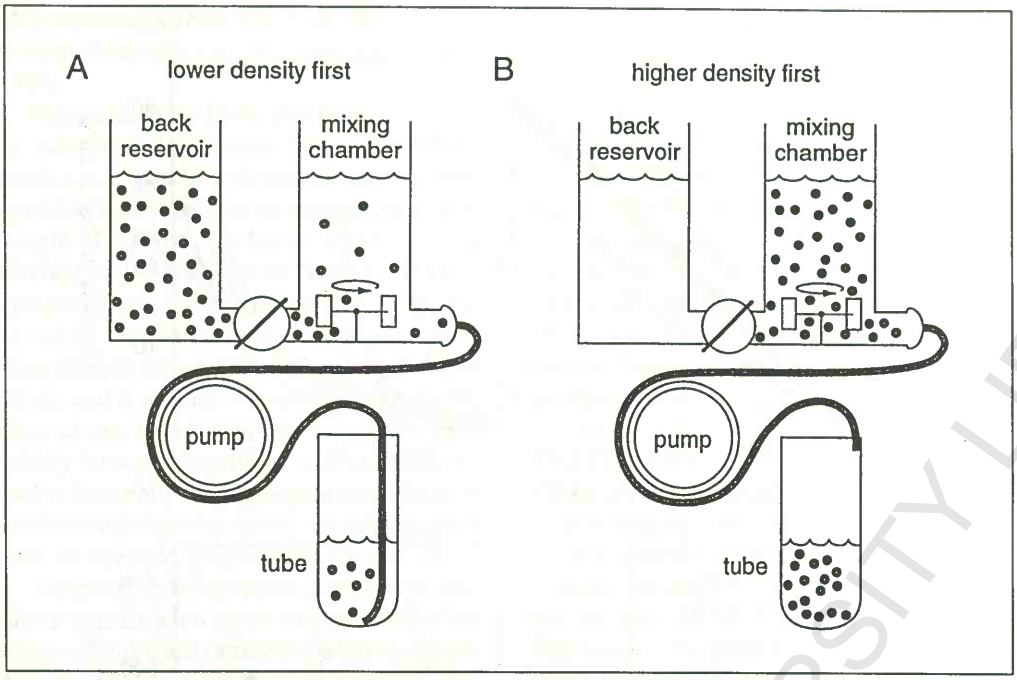


Figure 4.1.2 Linear gradient-forming device showing the back reservoir connected to the mixing chamber via a channel with a stopcock. A delivery tube leads from the mixing chamber to a peristaltic pump and then to the centrifuge tube. (**A**) Configuration used when the lower-density solution is delivered first; the lower-density solution is in the mixing chamber and the outlet is at the bottom of the centrifuge tube. (**B**) Configuration used when the higher-density solution is delivered first; the higher-density solution is in the mixing chamber and the outlet tube is always above the surface of the liquid entering the tube.

one chamber that connects to a peristaltic pump on its way to the centrifuge tube (see Fig. 4.1.2). The channel is gated by a stopcock. For linear gradients (the type most frequently used), the chambers have the same cross-sectional area. Solutions representing the extremes in density of the desired gradient are each placed in one chamber, and a stirring device is placed in the chamber with the outlet going to the tube. As fluid is pumped out of the mixing chamber, fluid also flows through the connecting channel from the reservoir to maintain the same pressure head in the two chambers. Consequently, the density of the solution in the mixing chamber progressively changes. Gradients can be generated either highest density first, if the higher-density solution is placed in the mixing chamber, or lowest density first, if the lower-density solution is placed in the mixing chamber. Note that for highest density first, the delivery point must be located above the level of the solution filling the centrifuge tube, whereas for lowest density first, the delivery point must be at the bottom of the tube.

Gradient-collection devices. Gradient collection at the end of centrifugation requires a systematic and reproducible means of emptying the tube. Gradients can be collected from the top or from the bottom. With care, a pipettor (e.g., Pipetman, Rainin) of suitable volume can be used for manual collection from the top, taking care to withdraw the equal-volume fractions from just below the meniscus. Alternatively, Buchler markets a pumping device (Autodensiflow) that continually adjusts to the height of the meniscus and withdraws gradients from the top, and both MSE Scientific and Nyegaard market devices that pump high-density solution beneath the gradient and progressively displace the gradient, which is again collected at the top. For collection from the bottom, the centrifuge tube is punctured with a needle and the gradient is collected dropwise.

FRACTIONATION MEDIA

Early in the development of centrifugation as a tool for purification of cellular organelles, it was established that nonelectrolyte-based media are preferable to electrolyte-based media for reducing organelle aggregation (Hogeboom et al., 1948). Most media used today are nonelectrolyte-based, although some do have a substantial electrolyte content (e.g., STKM medium, which contains 50 mM

BU MEDICAL CENTRAL UNIVERSITY LIBRARY

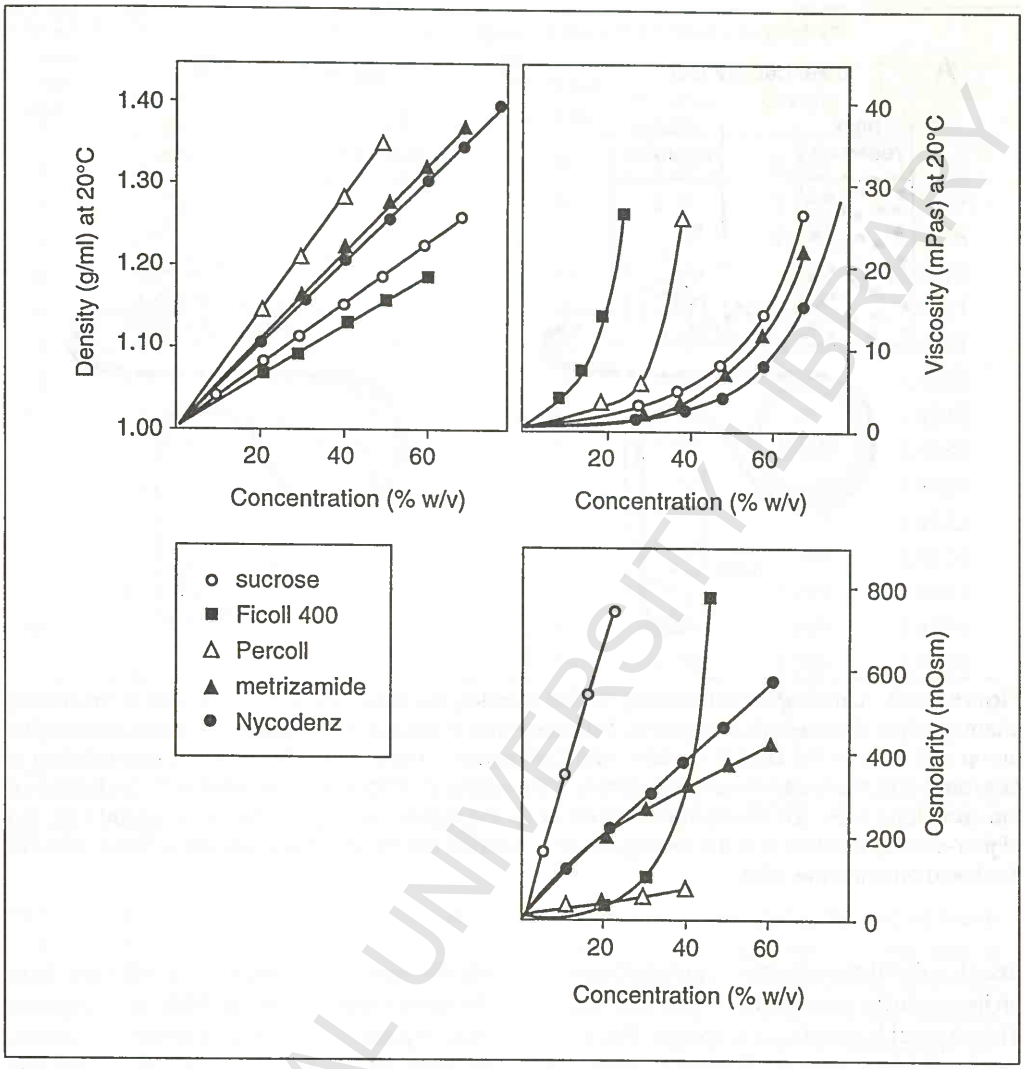


Figure 4.1.3 Plots of density, viscosity, and osmolarity of gradient media used to fractionate cellular organelles as a function of concentration. Data are shown for sucrose, Ficoll 400, Percoll, metrizamide, and Nycodenz. Modified from Rickwood (1984) by permission of IRL Press.

Tris·Cl, 25 mM KCl, and 5 mM $MgCl_2$ in addition to 0.25 M sucrose; Adelman et al., 1973). Moreover, media are generally isoosmotic or hyperosmotic as compared to 0.15 M NaCl. Organelles are osmometers, with water moving across their limiting membranes to maintain osmotic equilibrium; thus lysis can occur in hypoosmotic media.

The nonelectrolyte solute most commonly used in subcellular fractionation is sucrose. High-quality sucrose (free of RNase) is reasonably inexpensive, very soluble, and can be used readily to prepare solutions that span the range of densities of most biological organelles (Table 4.1.1). In addition, procedures have been developed for purifying most biological organelles in sucrose. Sucrose, however, has limitations; at higher concentrations solutions are quite viscous and hyperosmotic. Isoosmotic sucrose (0.25 M; ~9% w/v) has a density of only 1.03 g/ml, which is at the lower-density end for biological organelles. Thus all isopycnic density gradient centrifugation performed in sucrose constitutes hyperosmotic conditions; intraorganellar water redistributes to the surrounding medium, and the organelles shrink and increase in density. These changes are reversible for some, but not all, organelles.

Various sugar alcohols have been used in place of sucrose in some procedures. Glycerol is one, but although its viscosity is lower than that of sucrose, so is its density at a given concentration, and it also permeates some organelles. Mannitol has been used in a few procedures where the use of sucrose is not feasible—e.g., for isolation of epithelial brush border membranes that contain sucrase (Malathi et al., 1979)—and sorbitol has been a favorite

PCU IASRI CENTRAL UNIVERSITY LIBRARY

for fractionating yeast, which secrete invertase (a form of sucrase; e.g., Walworth and Novick, 1987).

Polysaccharides have also been employed as substitutes for sucrose. Ficoll 400 (Pharmacia), a chemical polymer of sucrose with epichlorohydrin that has an average molecular weight of 400,000, has been a popular choice (70,000 is also available, as Ficoll 70). At low concentrations, it has very low osmolarity, but as can be seen in Figure 4.1.3, the osmolarity rises sharply above 30% (w/v; density ~1.09 g/ml), and it is quite viscous at all concentrations above 10% (w/v). Ficoll has been used widely for separating different cell populations and is frequently most useful as an additive to media containing other density-modifying agents (see, for example, the procedures in *UNIT 4.2*).

Iodinated nonelectrolytes increase the density of fractionation media through most of the range of biological organelles while maintaining reasonably low viscosity and osmolarity as compared to sucrose. Two popular and closely related compounds are metrizamide and Nycodenz (both marketed by Nyegaard; Accurate Chemical—listed in the *SUPPLIERS APPENDIX*—is the U.S. outlet). The structures and specific application of these reagents are discussed in more detail, along with the procedure for their use, in *UNIT 4.2*. As seen from Figure 4.1.3, 40% (w/v) solutions of both metrizamide and Nycodenz have a density >1.2 g/ml (above that of most organelles) and an osmolarity ≤400 mOsm with reasonable viscosity. With all these advantageous features, why aren't they used more? They are rather expensive, which certainly limits their use for large-scale work. In addition, they absorb in the UV and have been reported to inhibit selected enzyme activities (see *UNIT 4.2*).

A different alternative for achieving high densities at low viscosity and osmolarity involves the use of Percoll, colloidal silica coated with polyvinylpyrrolidone to reduce its adsorption to biological organelles. For fractionation of cellular organelles, Percoll is almost always used in combination with isoosmotic sucrose, because pure Percoll is only 10 mOsm. As can be seen from Figure 4.1.3, Percoll-containing media can achieve very high densities with reasonably low viscosities. Because Percoll particles have a significant size (range 30 to 150 Å) and density, they sediment during centrifugation. Thus, concentration gradients self-form fairly rapidly during a spin and the profile of density along the length of the tube continuously changes with time (see Fig 4.2.2). As pointed out in *UNIT 4.2*, the shallow gradients achieved with Percoll can be very advantageous in separating organelles having densities that differ only slightly; however, for reproducibility, the conditions of centrifugation must be quite carefully set. Percoll also has its limitations: it absorbs in the UV, interferes with protein assays, precipitates in acid and organic solvents [do *not* attempt to precipitate Percoll-containing fractions with trichloroacetic acid (TCA) or acetone], and is difficult to get rid of entirely. Some strategies for removing Percoll are discussed in *UNIT 4.2*.

EVALUATION OF FRACTIONATION

In setting up a procedure for separating or purifying cellular organelles, it is advisable to evaluate the steps of the protocol to confirm that the outcome is as desired and as stated. This means setting up assays for marker activities and protein and accounting for the totality of both throughout the procedure. This meticulousness enables the investigator to evaluate the homogenization, yield, fold purification, and extent of contamination by other undesirable organelles, and to account for 100% of the activity at each step. The extent of cell breakage during homogenization can be estimated by determining the fraction of a marker activity of a particular organelle that pellets in the first low-speed centrifugation step (and is likely still to be associated with unbroken cells), while the extent of organelle breakage can be estimated by determining the fraction of the total amount of a soluble intraorganellar protein that is not sedimented by a high-speed spin that pellets all of the organelle. Following marker activity throughout purification enables the investigator to determine the yield of the desired organelle, and, if markers of other organelles are followed, the extent of residual contamination. Finally, if protein concentration is determined in parallel, calculation of the ratio of activity to protein (known as specific activity) at each step provides an estimate of fold purification or enrichment. In the event that analytical rather than preparative procedures are used, determination of marker activity in every fraction is essential for specifying the distribution of a particular organelle. Finally, it is very important to note that subcellular fractionation is never perfect! The entirety of a particular organelle is never obtained in completely pure form. In fact, the term "fraction" signifies both incomplete yield and incomplete purity. The careful investigator will always be cognizant

BUCKNELL UNIVERSITY LIBRARY

Table 4.1.2 Preparing Organelle Fractions from Mammalian Tissues and Cells

Organelle	Procedure	Reference
Nucleus	Centrifugation through high-density sucrose	Blobel and Potter (1966)
Endoplasmic reticulum	Discontinuous sucrose gradient	Adelman et al. (1973)
Golgi complex	Sucrose gradient, continuous[a] and discontinuous	Bergeron et al. (1982)
Secretion granules		
Endocrine	Metrizamide gradient	Loh et al. (1984)
Exocrine	Discontinuous sucrose gradient	Cameron and Castle (1984)
Synaptic vesicles	Chromatography on controlled-pore glass	Carlson et al. (1978) Huttner et al. (1983)
Plasma membrane	Discontinuous sucrose gradient	Hubbard et al. (1983)
Endosomes	Free flow electrophoresis Density shift with sucrose gradient[a]	Marsh et al. (1987) Beaumelle et al. (1990)
Lysosomes	Metrizamide gradient[a]	Wattiaux et al. (1978)
Mitochondria	Velocity sedimentation in sucrose	Schnaitman and Greenawalt (1968)
Peroxisomes	Sucrose gradient, discontinuous and continuous	Leighton et al. (1968)

[a]Procedure also presented in detail in *UNIT 4.2*.

of the limitations of the procedures and the consequent limitations of the results.

DEFINITIVE PROCEDURES

A useful starting point in isolating a particular cellular organelle may be to examine a procedure that has been designed for the specific organelle of interest. Table 4.1.2 lists references to procedures for purifying most of the membranous organelles found in eukaryotic cells (several are also presented as protocols in *UNIT 4.2*). Unfortunately, it is not possible to make this listing comprehensive, but the procedures selected have achieved unusually good purity and in many cases have documented their achievement by bookkeeping of marker activities. Although several of the procedures are reasonably old, they have been used either as starting points for subfractionation (e.g., nucleus, endoplasmic reticulum, secretion granules, plasma membranes, lysosomes, mitochondria, and peroxisomes) or for functional studies in cell-free assays (e.g., Golgi, endoplasmic reticulum, and mitochondria).

Literature Cited

Adelman, M.R., Blobel, G., and Sabatini, D.D. 1973. An improved cell fractionation procedure for the preparation of rat liver membrane-bound ribosomes. *J. Cell Biol.* 56:191-205.

Beaumelle, B.D., Gibson, A., and Hopkins, C.R. 1990. Isolation and preliminary characterization of the major membrane boundaries of the endocytic pathway in lymphocytes. *J. Cell Biol.* 111:1811-1823.

Bergeron, J.J.M., Rachubinski, R.A., Sikstrom, R.A., Posner, B.I., and Paiement, J. 1982. Galactose transfer to endogenous acceptors within Golgi fractions of rat liver. *J. Cell Biol.* 92:139-146.

Blobel, G. and Potter, V.R. 1966. Nuclei from rat liver: Isolation method that combines purity with high yield. *Science* 154:1662-1665.

Cameron, R.S. and Castle, J.D. 1984. Isolation and compositional analysis of secretion granules and their membrane subfraction from the rat parotid gland. *J. Memb. Biol.* 79:127-144.

Carlson, S.S., Wagner, J.A., and Kelly, R.B. 1978. Purification of synaptic vesicles from elasmobranch electric organ and the use of biophysical criteria to demonstrate their purity. *Biochemistry* 17:1188-1199.

Hofmann, G. (ed.) 1977. ISCO Tables, 7th ed. ISCO, Lincoln, Neb.

PCU IASTOMEDICAL UNIVERSITY LIBRARY

Hogeboom, G.H., Schneider, W.C., and Palade, G.E. 1948. Cytochemical studies of mammalian tissues. I. Isolation of intact mitochondria from rat liver; some biochemical properties of mitochondria and submicroscopic particular material. *J. Biol. Chem.* 172:619-635.

Hubbard, A.L., Wall, D.A., and Ma, A.K. 1983. Isolation of rat hepatocyte plasma membranes: I. Presence of the three major domains. *J. Cell Biol.* 96:217-229.

Huttner, W.B., Schiebler, W., Greengard, P., and DeCamilli, P. 1983. Synapsin I (Protein I), a nerve terminal–specific phosphoprotein: III. Its association with synaptic vesicles studied in a highly purified synaptic vesicle preparation. *J. Cell Biol.* 96:1374-1388.

Leighton, F., Poole, B., Beaufay, H., Baudhuin, P., Coffey, J.W., Fowler, S., and DeDuve, C. 1968. The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with Triton WR-1339. *J. Cell Biol.* 37:482-513.

Loh, Y.P., Tam, W.W.H., and Russell, J.T. 1984. Measurement of pH and membrane potential in secretory vesicles isolated from bovine pituitary intermediate lobe. *J. Biol. Chem.* 259:8238-8245.

Malathi, P., Preiser, H., Fairclough, P., Mallett, P., and Crane, R.K. 1979. A rapid method for the isolation of kidney brush border membranes. *Biochim. Biophys. Acta* 554:259-263.

Marsh, M., Schmid, S., Kern, H., Harms, E., Male, P., Mellman, I., and Helenius, A. 1987. Rapid analytical and preparative isolation of functional endosomes by free flow electrophoresis. *J. Cell Biol.* 104:875-886.

Rickwood, D. 1984. Centrifugation, A Practical Approach, 2nd ed. IRL Press, Oxford.

Schnaitman, C. and Greenawalt, J.W. 1968. Enzymatic properties of the inner and outer membranes of rat liver mitochondria. *J. Cell Biol.* 38:158-175.

Walworth, N.C. and Novick, P.J. 1987. Purification and characterization of constitutive secretory vesicles from yeast. *J. Cell Biol.* 195:163-174.

Wattiaux, R., Wattiaux-DeConnick, S., Ronveaux-Dupal, M.-F., and Dubois, F. 1978. Isolation of rat liver lysosomes by isopycnic centrifugation in a metrizamide gradient. *J. Cell Biol.* 78:349-368.

Contributed by J. David Castle
University of Virginia
Charlottesville, Virginia

B.C.U. „M. Eminescu" IAŞI
CHIMIE

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Purification of Organelles from Mammalian Cells

The protocols in this unit illustrate a range of different procedures that have been used to fractionate tissue homogenates. They emphasize different fractionation techniques that have been used for rat liver, an abundant tissue that has been a favorite of many investigators and has served as the source of many organelle preparations of excellent purity. For selected procedures, other examples have been given using other tissue sources (e.g., glandular tissues that maintain protein storage granules for regulated secretion) or, where particularly favorable, cultured cells.

The basis (or goal) of each separation and the merits and limitations of each procedure are summarized in Strategic Planning; this section serves as a guide for selecting among the various approaches (see Background Information for more in-depth discussion). In Strategic Planning the methods are described in somewhat general terms, but each corresponding protocol includes specific details relevant to a particular sample cell type (and in many cases a particular organelle).

STRATEGIC PLANNING

Because there are so many different techniques available for purifying organelles, the following annotated outline has been prepared as a guide to selecting a purification strategy.

Differential Centrifugation by Velocity (see Basic Protocol 1)

Basis/Goal

Enrichment of organelles largely according to size by serial spins of increasingly higher force and longer time.

Merits

1. This is the nearly universal starting point in fractionation of a homogenate to remove unbroken cells and large debris (as well as nuclei, if not desired).
2. The procedure is the best way to concentrate organelles of a particular size range from large volumes for use as starting material for subsequent procedures.
3. It is generally performed in isoosmotic sucrose medium, thus providing the best chance of organelles retaining function in subsequent cell-free assays.
4. It is useful for sedimenting macromolecular assemblies, e.g., microtubules polymerized from organelle-free postmicrosomal supernatant.
5. The separation is simple to perform. Within limits, successive resuspension and recentrifugation can be used to further enrich a pellet for a given size range of organelle.

Limitations

1. Pelleted fractions are nearly always substantially contaminated with smaller organelles.
2. Aggregation of organelles in pellets can be significant and difficult to reverse, especially if medium has a substantial electrolyte content.

Copyright © 1995 by John Wiley & Sons, Inc.

BOU IAS / CENTRAL UNIVERSITY LIBRARY

Differential Centrifugation by Equilibrium (see Basic Protocol 2)

Basis/Goal
Separation of organelles according their distinct buoyant densities by centrifuging to equilibrium in a density gradient.

Merits
1. The technique provides an excellent follow-up to differential centrifugation by velocity (Basic Protocol 1) for resolving similar-sized organelles into subpopulations of distinct densities.
2. Use of this procedure is often advantageous for complex mixtures resolved as multiple bands.
3. It is also useful for analytical experiments in which the distribution of organelles (or co-distribution of proteins with particular organelles) is being examined, and for preparative experiments to obtain purified organelles for further study.

Limitations
1. Extended centrifugation times may be detrimental to labile activities.
2. Quantities of sample loaded on each gradient must usually be low to maintain high resolution and avoid aggregation.

Rate Zonal Centrifugation Using Sucrose (see Alternate Protocol 1)

Basis/Goal
Resolution of organelles from a loading layer (most often at the top of the tube) into enriched zones within a density gradient; separation based on size, shape), and density.

Merits
1. The procedure provides multiparameter separation within a single run, which is especially advantageous for analytical experiments that rely on different distributions for several organelles.
2. It is widely used for separations of macromolecules and macromolecular assemblies in addition to membrane-bounded organelles: e.g., proteins of different molecular weights (either water-soluble proteins or detergent-solubilized membrane proteins); polysomes of different length.

Limitations
1. Organelle/protein content of the sample loaded must be kept reasonably low for best resolution; aggregation may be a problem with overly large loads.
2. Its main value is in analysis of organelle distribution; use in organelle purification generally requires further steps.
3. Hyperosmotic high-density sucrose solutions may cause organelle shrinkage due to water efflux, and can thereby reduce density differences between certain types of organelles.

Gradient Fractionation Using Percoll (see Alternate Protocol 2)

Basis/Goal
Rapid separation largely by density in an isoosmotic self-forming density gradient.

CU IAST CENTRAL UNIVERSITY LIBRARY

Merits

1. Percoll's negligible osmolarity but high density and modest viscosity allow its use as an additive to sucrose for achieving a very wide range of densities under isoosmotic conditions.
2. Because Percoll itself sediments appreciably during centrifugation, self-formed density gradients continually change shape during the run. Initially, they are very shallow along most of the length of the tube, allowing excellent separation of organelles that have only slightly different densities and that approach isopycnic density rather rapidly. Large organelles with slightly differing densities can be easily resolved in fixed-angle rotors, whereas small organelles with slightly differing densities can be resolved in vertical rotors (where the short path length during the run reduces the time needed to achieve equilibrium).
3. Equilibrium density separations require relatively short centrifugations, especially compared to those in high-density sucrose.
4. The procedure can be used effectively for large-scale organelle purification, especially where the organelle of interest has either a lower or a higher density than all other organelles in the starting mixture.

Limitations

1. In general, Percoll is not very suitable for generating distinct distributions for several different organelles in the same centrifuge run. With a gradient profile featuring a shallow central segment between steep extremes, low-density organelles tend to concentrate together at the top of the gradient while high-density organelles tend to concentrate together at the bottom of the gradient.
2. Thoroughly removing Percoll from a purified fraction is a major challenge (see Critical Parameters for further discussion).
3. Further limitations may be imposed by the organelles themselves. Certain organelles may have heterogeneous sizes, shapes, and densities; thus, portions of one type of organelle may end up at both extremes of the gradient.

Gradient Fractionation Using D_2O (see Alternate Protocol 3)

Basis/Goal

Separation by density gradient with heavy (deuterated) water substituted for normal water.

Merits

1. There is reduced osmolarity and viscosity of fractionation medium.
2. It is mainly used in analytical procedures.

Limitations

1. The procedure is quite expensive due to the cost of D_2O, limiting its preparative use.
2. Partial exchange of D_2O for H_2O occurs within membrane-bounded organelles, with the extent of exchange differing among organelles. This may either enhance or reduce their separation and requires empirical analysis.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Gradient Fractionation Using Nonionic Iodinated Solutes (see Alternate Protocol 4)

Basis/Goal

Separation by density gradient with iodinated nonelectrolytes substituted for sucrose to achieve high density.

Merits

Substituting nonionic solutes for sucrose makes it possible to prepare media that span the range of densities of all membrane-bounded organelles, yet have quite low viscosities and only moderate osmolarities. For organelles with a high buoyant density, this significantly reduces efflux of water, which can lead to irreversible changes in some organelles.

Limitations

1. The biggest limitation to these very promising media—metrizamide and Nycodenz—is their expense, which prevents most labs from using them for routine large-scale fractionation.
2. Selected enzymes are inhibited by these media but most are not; see Alternate Protocol 4 (explanatory Note preceding the protocol).
3. It is important to keep in mind that these media absorb in the UV range.

Gel Filtration to Isolate Secretory Vesicles (see Basic Protocol 3)

Basis/Goal

Purification of small, homogeneous vesicles from larger and/or irregularly shaped organelles.

Merits

1. The procedure offers excellent resolution for purifying some of the smallest membrane-bounded organelles.
2. It is a promising analytical technique for demonstrating specific association of selected proteins with small vesicles.
3. It is very useful for characterizing or purifying protein complexes, e.g., the cytosolic macromolecular assemblies that function in membrane coat formation during vesicular trafficking.

Limitations

1. Separation is limited to organelles that are small enough to enter the pores of the gel filtration resin efficiently. As indicated by examples cited in the protocol, this means vesicles <200 nm diameter.
2. As the separation is based on size alone, gel filtration with the available resins will not distinguish different types of small vesicles having the same size.

Preparative and Analytical Electrophoretic Separation (see Basic Protocol 4 and Alternate Protocol 5)

Basis/Goal

Separation of small organelles and macromolecular assemblies on the basis of distinct surface charge densities.

Merits

1. Although electrophoretic separations have not been used widely in organelle purification, the preparative technique (Basic Protocol 4) shows substantial promise

CU IAS / CENTRAL UNIVERSITY LIBRARY

because it is reasonably easy, not expensive, and as reported by its originators (L. Rome, pers. comm.) is capable of handling relatively large samples (i.e., 50 mg protein). Furthermore, preliminary findings suggest that larger-porosity agarose gels might be applicable in the separation of organelles up to 300 nm in diameter.

2. The analytical procedure (Alternate Protocol 5) requires a free-flow electrophoresis chamber, either commercial or constructed as described by Tulp et al. (1993); the latter can be done for only a minor fraction of the cost of the commercial apparatus. The impressive separations that have been reported (Tulp et al., 1994) make this device worthy of serious consideration. Alternatively, electrophoretic analyses of synaptic vesicles (Carlson and Kelly, 1978) have been performed on Ficoll gradients in plastic tissue culture pipets.

Limitations

Use of these procedures has been limited to work on coated vesicles or endosomes/lysosomes; therefore, their applicability would need to be tested before they could be applied to other organelles.

Solid-Phase Immunoadsorption (see Basic Protocol 5)

Basis/Goal

Isolation of vesicles or macromolecular assemblies using specific antibodies attached to solid-phase supports.

Merits

1. The technique is very promising and highly selective.
2. It can be used both analytically and preparatively. Solid phases include fixed *S. aureus*, protein A–Sepharose, and secondary or primary antibody coupled to several types of beads (e.g., Sepharose, Eupergit, magnetic) or other polymers.
3. Co-localization of components on immunoadsorbed vesicles can be demonstrated secondarily by immunochemical or immunocytochemical procedures (e.g., see Laurie et al., 1993).

Limitations

1. Immunoadsorption requires a specific antibody that can bind to an exposed epitope on the surface of the organelle (or macromolecular assembly). The success of the technique depends on the quality of the antibody (specificity and avidity) and accessibility of its epitope; adsorption of vesicles derived from intracellular compartments requires cytosolically exposed epitopes, whereas adsorption of plasma membrane vesicles (which generally vesiculate with the ectodomain outward) requires epitopes on the extracellular surface.
2. Although sensitivity is generally a positive quality, in immunoadsorption it can be a limitation. Vesicle binding via an avid antibody may require only a few antigen-antibody interactions. As many membrane proteins are not confined to single compartments but instead traffic among compartments, immunoadsorption may not necessarily isolate only a single type of compartment.

Purification by Lectin Adsorption (see Basic Protocol 6)

Basis/Goal

Isolation of vesicles derived from plasma membrane by lectin binding to exposed oligosaccharides.

SCU IRSC / CENTRAL UNIVERSITY LIBRARY

Merits

1. The procedure takes advantage of the unique orientation (ectodomain outward) of plasma membrane vesicles. Exposed glycoproteins on these vesicles can bind to immobilized lectins, particularly wheat germ agglutinin (WGA).
2. The procedure is organelle-specific and quite simple as there is limited need for preenrichment steps.
3. *N*-acetylglucosamine used to elute vesicles from WGA is inexpensive.

Limitations

1. The procedure is largely specific for a single type of membrane.
2. Separation is only as good as the uniformity of orientation of vesiculated membranes. Intracellular vesicles with inverted orientation can potentially bind to WGA, and inverted plasma membrane vesicles will be unable to do so. Respectively, these occurrences will increase contamination and reduce yield. Generally, however, the orientations are fairly uniform, so this does not pose a major problem.
3. Although the purification of plasma membrane is excellent using the protocol presented here, in other procedures it has been observed to be less extensive; possibly adsorbed vesicles or cytoskeletal proteins increase the contamination.

Density Shift Using Digitonin (see Basic Protocol 7)

Basis/Goal

Analytical procedure in which binding of digitonin to cholesterol increases the buoyant density in proportion to the cholesterol content of the membrane. Density shifts determined following equilibrium centrifugation range from zero (for ER and mitochondria, which have no resident cholesterol) to intermediate (Golgi cisternae) to maximal (post-Golgi membranes, including the plasma membrane).

Merits

An excellent procedure for distinguishing the possible association of a newly identified (or newly synthesized) protein with a particular compartment, using the extent of density shift of compartmental markers for correlation.

Limitations

Correlation of distributions is a good supporting technique but does not directly demonstrate a particular association.

Density Shift Using Colloidal Gold Conjugates (see Basic Protocol 8)

Basis/Goal

Capitalizes on the high density of colloidal gold and the well-established characteristics of receptor-mediated endocytosis to provide density modification of selected endocytic compartments, thereby enabling their purification.

Merits

1. Although this procedure has not been used widely, it appears quite promising for studies aimed at characterizing the various compartments comprising the endocytic network.
2. The density shifts observed are substantial, making this a good preparative method,

ACU IAS / CENTRAL UNIVERSITY LIBRARY

and the procedure can be tailored to specific pathways based on the choice of the ligand that is conjugated to colloidal gold.

Limitations

1. Preparation of the gold conjugates is fairly lengthy, but the procedures involved are not difficult.
2. Procedures are limited to cells or tissue where it is possible to introduce endocytic tracers efficiently.
3. The purity of the fractions obtained is limited by the ability to achieve selective compartmental labeling. No endocytic tracers are perfect compartment markers; thus density shifts are likely to spill over into sequential compartments.

Extraction of Extrinsic Proteins from Membranes Using Sodium Carbonate (see Basic Protocol 9)

Basic/Goal

Stripping of extrinsic (or adsorbed) proteins from membrane vesicles in order to distinguish candidate integral membrane proteins as those that remain membrane bound.

Merits

1. Widely used analytical tool to distinguish membrane proteins anchored in the lipid bilayer either by transmembrane segments or by lipid tethers (e.g., prenylation).
2. Frequently a complementary procedure to phase partitioning in the nonionic detergent Triton X-114 (Bordier, 1981).

Limitations

1. The extreme pH may denature proteins of interest; this possibility should be tested on an individual basis.
2. Membrane vesicles develop discontinuities during treatment (see Fujiki et al., 1982), so subsequent procedures should not require intact vesicles.

DIFFERENTIAL CENTRIFUGATION BY VELOCITY

BASIC PROTOCOL 1

This is the classical standard method of cell fractionation using centrifugation, often used as the starting point for a more complex purification procedure. A cell homogenate is subjected to serial centrifugations of increasingly higher force and longer duration, generally in isoosmotic (0.25 to 0.3 M) sucrose medium (see Background Information for discussion of sucrose concentration). Resulting gradient bands (concentrated fractions consisting of organelles of particular sizes) are suitable as starting material for further purification or other procedures. This method gives relatively low resolution in that it does not yield fractions of high purity. However, the postnuclear supernatant (PNS) from step 6 and the supernatant from step 8 serve as useful starting materials for procedures that yield high-purity fractions mainly by exploiting differences in organelle density.

As an example of the approach, this protocol presents the fractionation of rat liver into various components, including mitochondria, microsomes, and a largely organelle-free "cytosol" fraction. Additional steps for using tissues other than liver are included.

BCU ARTS CENTRAL UNIVERSITY LIBRARY

Materials

Rat starved overnight and freshly sacrificed
Homogenization medium: 0.25 to 0.3 M sucrose plus optional additives (see recipe; if PMSF is included, add just before use)

Motor-driven pestle with high-torque motor (e.g., Baxter)
Cheesecloth
Tubes suitable for low-speed centrifugation: 1.5-ml microcentrifuge tubes (for small-scale preparation) or 15- or 50-ml graduated screw-cap tubes (for larger-scale preparations)
Low-speed centrifuge capable of producing a 600 × *g* centrifugal field, 4°C
Teflon-glass homogenizer (A.H. Thomas or equivalent) of appropriate size for the scale of the experiment: type AA (1 ml) up to type C (40 ml)
200-µm mesh nylon screen (Tetko)
Dounce-type glass homogenizer
Polycarbonate or polyallomer tubes suitable for ultracentrifugation (e.g., 3-ml size to fit Beckman TLA-100.3 rotor) or polycarbonate bottle assemblies (e.g., 10-ml size to fit Beckman 50Ti rotor or 25-ml size to fit Beckman 60Ti or 70Ti rotor)
Ultracentrifuge and fixed-angle rotor appropriate to the scale of the experiment (e.g., Beckman 50Ti, 60Ti, or 70Ti) or tabletop ultracentrifuge and rotor (e.g., Beckman tabletop with TL100 rotor)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Homogenize the sample

1. Remove liver from a freshly sacrificed rat that has been fasted overnight (to deplete hepatic glycogen). Keeping the liver on ice as much as possible, weigh and then mince coarsely with scissors or razor blade in the weighing boat.

   Stored frozen tissues are not satisfactory for most organelle purification procedures.

2. Homogenize the minced liver at 20% (w/v) in homogenization medium using five up-down strokes at 1700 rpm with a high-torque motor-driven pestle.

3. Filter the homogenate through four layers of cheesecloth.

Perform low-speed centrifugation to remove nuclei and larger particulates

4. Transfer the homogenate to tubes suitable for low-speed centrifugation and centrifuge 10 min at 600 × *g* (1700 rpm for a 20-cm-radius rotor), 4°C.

   For small-scale procedures, this centrifugation can be performed in 1.5-ml microcentrifuge tubes for 40 sec at 10,000 rpm (6600 × g).

5. *For tissues such as pancreas and salivary gland:* Resuspend the pellet in the original volume of homogenization medium and rehomogenize with the Teflon-glass pestle, then centrifuge again under the same conditions. Combine the resulting supernatant with the supernatant from the first centrifugation, filter through 200-µm mesh nylon screen, and disperse by 3 to 4 strokes in a Dounce homogenizer.

   Pancreas and salivary gland (and some other tissues) are more difficult to homogenize than liver. Homogenization is one of the key steps in successful organelle purification procedures, so it is important to employ the two-step homogenization procedure when appropriate.

6. The pellet is the crude nuclear fraction and the supernatant (or combined supernatants,

UCU IAS / CENTRAL UNIVERSITY LIBRARY

for optional two-step homogenization) is the postnuclear supernatant (PNS) fraction. Aspirate any lipid that has accumulated at the surface of the PNS, then withdraw the PNS, taking care to leave the pellet undisturbed. Transfer the PNS to polycarbonate tubes that are suitable for ultracentrifugation.

Prepare mitochondrial and microsomal fractions by successive ultracentrifugations

7. Centrifuge the PNS 30 min at 10,000 × *g* (e.g., 10,000 rpm in Beckman 50Ti, 60Ti, or 70Ti), 4°C, to pellet the mitochondrial fraction.
8. Remove the supernatant and transfer to clean polycarbonate ultracentrifuge tubes.
9. Resuspend the mitochondrian-rich pellet and transfer to a Dounce homogenizer. Disperse the pellet into a fine suspension by several manual strokes.

   The pellet also contains lysosomes and peroxisomes (which are much less abundant) and small amounts of rough and smooth microsomes that have happened to pellet.

10. Optionally, further purify the mitochondria away from smaller-sized contaminants by repeating steps 6 and 7.

    This strategy has been used by Carvalho-Guerra (1974) for preparing liver mitochondrial fractions.

11. Centrifuge the supernatant from step 7 for 2 hr at 100,000 × *g*, 4°C, in fixed-angle rotor.
12. The resulting pellet is the total microsomal fraction, consisting of vesicles derived from rough endoplasmic reticulum and various smooth-membrane-bounded organelles (including fragments from mitochondria and larger-sized organelles that were damaged during homogenization). Remove the postmicrosomal ("high-speed") supernatant; this is generally used as "cytosol" in various assays.
13. Resuspend the microsomal pellet in an appropriate buffer for assays or additional purification, disperse by Dounce homogenization, and save at 4°C until further use.

DIFFERENTIAL CENTRIFUGATION BY EQUILIBRIUM

BASIC PROTOCOL 2

This procedure can be used alone (as presented here) but most often is used as a follow-up to differential centrifugation by velocity (Basic Protocol 1). Typically, the samples are applied above (or beneath) a sucrose or other nonelectrolyte-based medium having a density that varies—either discontinuously (in steps) or continuously—from top to bottom of the tube. The sample must of course be lower yet in density than the lowest density of the gradient (if applied to the top) or higher than the highest density of the gradient (if applied at the bottom). During centrifugation, organelles sediment (or float) until reaching their isopycnic positions, where the buoyant density of the particle equals the density of the medium (*UNIT 4.1*). In a continuous gradient organelles separate as bands distributed throughout the gradient; in a discontinuous gradient the organelles concentrate at the interfaces between the particular steps whose densities are lower and higher than that of each organelle. A continuous gradient is preferred for fractionating complex mixtures containing several different organelles.

As an example of this type of fractionation, the following one-step procedure (taken from Bergeron et al., 1982) describes separation of fractions from homogenates of rat liver by flotation in a continuous sucrose gradient. Specific information about collecting Golgi fractions (the goal of Bergeron's work) is described in annotations.

CHILE
BOSTON CENTRAL UNIVERSITY LIBRARY

Additional Materials (also see Basic Protocol 1)

Rat starved overnight and freshly sacrificed
Sucrose media: 0.25 M, 1.02 M, and 1.8 M sucrose (ultrapure; e.g., ICN Biochemicals) in dilution solution
Dilution solution: 50 mM Tris·Cl (pH 7.4)/25 mM KCl/5 mM $MgCl_2$
Ultraclear or polyallomer ultracentrifuge tubes to fit a swinging-bucket rotor (e.g., 38 ml for Beckman SW28 rotor or equivalent; 13 ml for SW41 rotor or equivalent; 5 ml for SW50.1 or SW55 rotor or equivalent)
Gradient-forming device (see Fig. 4.1.2)
Ultracentrifuge and swinging-bucket rotor (e.g., Beckman SW28, SW41, SW50.1, or SW55)
Syringe (with blunt needle), plastic transfer pipet, or siliconized glass Pasteur pipet

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare sample by homogenization and low-speed centrifugation

1. Remove liver from a freshly sacrificed rat that has been fasted overnight (to deplete hepatic glycogen). Weigh liver, mince, homogenize, and filter through cheesecloth as for velocity centrifugation (see Basic Protocol 1, steps 1 to 3), except homogenize at 15% (w/v) in 0.25 M sucrose medium.
2. Add 1 vol of 1.8 M sucrose medium to the homogenate to give 1.02 M sucrose final.

Load and run gradients

3. Load the adjusted homogenate into centrifuge tubes to fill ≤½ the tube volume.

   It is important not to overload individual gradients with sample, as overloading promotes aggregation and decreases resolution and purity.

4. Use a gradient-forming device to form a continuous sucrose gradient on top of the loaded samples. To form the gradient, place 0.25 M sucrose solution in the back reservoir of the gradient-forming device and 1.02 M sucrose solution in the mixing chamber. Open the connecting channel and pump the gradient into the tube (high density first), making sure that the outlet remains above the level of the solution in the filling tube (see Fig. 4.1.2).

   Alternatively, the continuous gradients can be prepared in advance and kept at 4°C for a few hours before loading. In this case, use a syringe with a long needle to carefully load the 1.02 M–adjusted homogenate beneath the gradient in each tube.

5. After the gradients are made, insert the tubes into the buckets of the swinging-bucket rotor. Weigh buckets plus tops and balance in pairs by adding a small amount of the lowest-density sucrose medium to the lighter tube of each pair.
6. Centrifuge tubes 3 hr at 100,000 × *g* (e.g., 28,000 rpm in Beckman SW28), 4°C. Allow the rotor to coast to a stop with the brake off.

   If the deceleration is too rapid, the layers of sucrose (e.g., between the load zone and the gradient) rotate relative to one another, which causes vortexing into the overlying sucrose layer(s) and decreases the resolution. With older ultracentrifuges (Beckman L5-50, L5-65, and older), the rotor coasts to a stop from maximal speed, taking ~30 min from 28,000 rpm. As most of the disturbance of gradients occurs below the final 1000 rpm of deceleration, newer instruments decelerate to 800 rpm and then coast to

CU IASI CENTRAL UNIVERSITY LIBRARY

a stop, taking significantly less time. Run times are given for older instruments, and can be adjusted downward to accommodate the new deceleration pattern.

Unload the gradients and collect fractions

7. Carefully remove the gradients from the rotor buckets, place in a rack, and note the positions of the bands.

   For the example described here (starved rat liver), a prominent band was found to have floated ~⅓ of the way into the gradient (i.e., ~⅓ of the way up) while a second band was observed at the air/0.25 M sucrose interface.

8. Begin collecting the fractions (bands) separately starting from the top, using a syringe (with blunt needle), plastic transfer pipet, or siliconized glass Pasteur pipet. Be sure to collect thoroughly: first sweep the circumference of the tube with the syringe or pipet tip, then pass the tip systematically across the interface while applying suction.

   Because the direction of centrifugation is radial, bands are more highly concentrated at the walls of the tube than in the center (see UNIT 4.1 for discussion of rotor geometry).

9. If recovery of total activity of the loaded sample is a goal of the experiment, collect aliquots of the solution between bands, as well as aliquots of the original sample and the pellet (after resuspension), as separate samples for assay.

Process individual fractions

10. Dilute each band at least 2- to 3-fold with buffer or a low-concentration sucrose medium. Pellet by velocity sedimentation for ≥100,000 × *g* (e.g., 40,000 rpm in Beckman SW41 or SW50.1), 4°C, then resuspend in the same medium for further use.

    Process Golgi fraction by adding dilution solution to 0.25 M sucrose final (use a refractometer, if available, to measure the sucrose concentration), then pelleting 1½ hr at 100,000 × g, 4°C.

    Generally bands obtained by isopycnic centrifugation are not as concentrated as resuspended pellets obtained by differential velocity sedimentation (see Basic Protocol 1). Furthermore, individual bands are suspended in media with differing densities.

11. If desired, subfractionate the samples further—for example, by lysis to separate membranes from soluble content or by high pH (carbonate) treatment, which is frequently used to strip membranes of extrinsic proteins (see Basic Protocol 9).

    As an alternative to the continuous gradient procedure, it is also possible to generate Golgi fractions (and especially subfractions enriched in different Golgi cisternae) by flotation on a discontinuous sucrose gradient. In this case, the starting sample is a crude microsomal pellet (see Basic Protocol 1) that has been resuspended in 1.15 M sucrose/50 mM Tris·Cl (pH 7.4)/25 mM KCl/5 mM $MgCl_2$. Above this sample in the centrifuge tube are placed layers of 1.0 M, 0.86 M, 0.6 M, and 0.25 M sucrose solution, and the tubes are spun as above. Bands are then collected at each interface: the 0.25 M/0.6 M interface is the light Golgi fraction; the 0.6 M/0.86 M interface is the intermediate Golgi fraction; the 0.86 M/1.0 M interface is the heavy Golgi fraction; and the 1.0 M/1.15 M interface is the small vesicular fraction (Bergeron et al., 1982).

    Because the liver produces very low density lipoproteins that transiently accumulate in the Golgi before export to the cell surface, Golgi from this source has an unusually low buoyant density and is especially favorable for preparing reasonably purified fractions. Golgi fractions from other sources are more dense, and so far most preparations from nonhepatic sources are not as pure.

BOSTON UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 1

RATE ZONAL CENTRIFUGATION USING SUCROSE

Rate zonal centrifugation is a nonequilibrium density gradient procedure for fractionating cells, in which sedimentation depends on the size, density, and shape of the particles as well as the viscosity and density of the medium. Typically, samples are loaded atop a continuous sucrose gradient; during subsequent reasonably short centrifugations (length of run and rate often need to be determined empirically), they separate as distinct bands (zones) enriched in particular organelles. This procedure can be adapted to various-sized gradients; available models of gradient-forming devices can be used to generate gradients in tubes ranging from 5 to 38 ml. To prepare gradients in smaller tubes (e.g., for a tabletop ultracentrifuge), investigators often approximate continuous gradients by preparing step gradients having several layers of progressively decreasing density, then letting them stand a few hours before use. Diffusion between the layers makes the approximation quite good.

The example described in this protocol is the separation of a postnuclear supernatant (PNS) from exocrine pancreas. It can be adapted to other tissues or cell cultures but this may require adjusting the density limits of the gradient. See, for example, the preparation of Golgi membranes for cell-free assays (Balch et al., 1984). Figure 4.2.1 presents the results of assaying marker activities of several organelles across the gradient. Ficoll 400 (a hydrophilic polymer; see *UNIT 4.1*) has been added to the sucrose to retard the migration of mitochondria relative to zymogen granules so that these two abundant organelles are resolved as distinct bands.

Materials

Rat starved overnight and freshly sacrificed
Homogenization medium: 0.3 M sucrose (ultrapure; e.g., ICN Biochemicals) plus 0.2 µg/ml diphenyl-*p*-phenylenediamine (DPPD, Kodak; from 0.4 mg/ml stock in ethanol)
1 M MES
0.2 M EDTA (*APPENDIX 2E*)
Gradient stock solutions, low- and high-density (see recipe)
0.3 M sucrose/5 mM MES/1 mM EDTA/0.2 µg/ml DPPD, pH 6.5 to 6.7

Cheesecloth
Dounce homogenizer
Gradient-forming device (see Fig. 4.1.2) suitable for selected gradient size, including compatible magnetic stir bars, tubing, and peristaltic pump
Ultraclear or polyallomer centrifuge tubes (of the selected size) suitable for swinging-bucket rotor
Plastic tissue culture pipet or polyethylene transfer pipet
Ultracentrifuge and appropriate swinging-bucket rotor (e.g., Beckman SW50.1 or SW55 if using 5-ml tubes; Beckman SW28 for 38-ml tubes)
Automated gradient-collection device (e.g., Buchler Autodensiflow, Nyegaard or MSE Scientific) *or* micropipettor for manual collection

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare the pancreatic samples

1. Remove pancreas from freshly sacrificed rat that has been fasted overnight (to maximize the zymogen granule population). Weigh pancreas (usually 0.8 to 1 g per animal), mince, homogenize, and filter through cheesecloth (as for liver; see Basic

CU IASL CENTRAL UNIVERSITY LIBRARY

Protocol 1, steps 1 to 3), except use homogenization medium containing 0.2 μg/ml DPPD and homogenize at 15% (w/v).

DPPD, an antioxidant, is added because it significantly reduces lysis of zymogen granules.

2. Centrifuge sample, rehomogenize (in the same volume as initially used of homogenization medium containing DPPD), and recentrifuge, then pool supernatants (see Basic Protocol 1, steps 4 and 5). Add MES to 5 mM (from 1 M stock) and EDTA to 1 mM (from 0.2 M stock). Filter and disperse with a Dounce homogenizer (see Basic Protocol 1, step 5). Maintain homogenized sample (PNS) at 0°C, preferably briefly, until loaded on the gradients.

Prepare the continuous gradients

3. Set up the gradient-forming device and peristaltic pump as shown in Fig. 4.1.2.
4. Load the gradient-forming solutions. Into the back chamber (reservoir) of the gradient former, pipet the high-density sucrose solution (15 ml for 38-ml tubes; 5 ml for 12.5-ml tubes; 2.2 ml for 5-ml tubes). Carefully bleed a little of the solution through to the front (mixing) chamber to displace the air in the joining tube. Then pipet the low-density solution into the mixing chamber (same volumes as for high-density solution). Be sure that in this configuration that the delivery tube is at the bottom of the centrifuge tube, as the gradient will be generated low density first (see Fig. 4.1.2).
5. Start the stirring motor and the peristaltic pump, and immediately open the joining tube between the two chambers. Note the Schlieren patterns in the mixing chamber as high-density sucrose flows in, first as is, then mixed. Set the delivery rate of the peristaltic pump empirically; if it is too fast, the solution level in the mixing chamber will drop below that in the delivery chamber and mixing will appear incomplete.
6. Once delivery of the solutions from the gradient-forming device to the centrifuge tubes is complete, shut off the peristaltic pump before air in the emptying delivery tube reaches the gradient. Slowly withdraw the delivery tube up through the gradient so disturbance to the gradient is minimal.

Continuous gradients are quite stable and can be prepared the day before use provided they are stored at 4°C without disturbance.

Load the sample on the gradients and centrifuge

7. Using a plastic tissue culture pipet or polyethylene transfer pipet, carefully overlay the gradients with samples of PNS. Maximal volumes of the 7.5% (w/v) PNS that can be loaded without overloading and decreasing resolution are 5.5 ml on a 30-ml gradient, 1.8 ml on a 10-ml gradient, and 0.5 ml on a 4.5-ml gradient.

To the inexperienced investigator, use of pipets may sound tricky. In our experience, nice layering with little mixing can be achieved rather quickly by placing the tip of the pipet on the meniscus of the solution while slightly tilting the tube containing the partly layered gradient, then gradually relaxing the pressure on the finger controlling the flow from the pipet. Some practice may be necessary! As an alternative, polyethylene transfer pipets can be used to load the solutions.

8. Place the loaded gradients into the buckets of the appropriate swinging-bucket rotor, taking care that the tubes are not wet on the outside. Balance opposite buckets to within 0.1 g of each other using homogenization medium.

Generally, with accurate pipetting during forming and loading, only minor adjustment is required for balancing.

9. Centrifuge the samples in the ultracentrifuge set at 4°C. For rate zonal centrifugation,

BODLEIAN LIBRARY OXFORD UNIVERSITY

which is a nonequilibrium procedure, the angular velocity and length of the run may require empirical adjustment.

For pancreatic PNS centrifuged in an SW28 rotor, excellent separation conditions are obtained by spinning 110 min at 26,000 rpm (90,000 × g), 4°C, then letting the rotor coast to a stop without the brake.

10. At the end of the run, the gradients should show discrete bands or zones that are enriched in specific organelles; these zones are distributed throughout the gradient.

Collect the gradients at 4°C

11. Unload the gradients in successive fractions of equal volume using an automated collection device or a micropipettor. For large-volume gradients in 38-ml tubes, a Buchler Autodensiflow device, which collects from the top, works very well. For small-volume gradients, collection can be satisfactorily performed using a micropipettor, taking care to withdraw the solution slowly and from as close to the surface as possible.

12. Assay the gradient fractions for the activity of interest. For some assays it may be necessary to reduce the sucrose content and/or concentrate the sample. If so, dilute fractions 2- to 3-fold with 0.3 M sucrose/5 mM MES/1 mM EDTA/0.2 µg/ml DPPD and pellet 1 hr (or shorter if the volume is small) at 100,000 × *g*, 4°C.

Figure 4.2.1 shows the results obtained by assaying the gradient fractions obtained by rate zonal centrifugation of the pancreatic PNS for enzymes that mark particular organelles.

ALTERNATE PROTOCOL 2

GRADIENT FRACTIONATION USING PERCOLL

Percoll (polyvinylpyrrolidone-coated colloidal silica) is a useful alternative to high-concentration sucrose gradients, particularly where high densities combined with low viscosity and osmotic activity are desired. Cells are homogenized, mixed with a Percoll solution, and centrifuged, generally using a fixed-angle rotor. The Percoll forms its own density gradient during centrifugation (a self-forming gradient). As it has very low osmotic activity (pure Percoll as purchased is ~10 mOsm), it is generally added to isoosmotic (or slightly hyperosmotic) sucrose for separating organelles from cell homogenates under isoosmotic (or near isoosmotic) conditions. The density range is controlled by the Percoll concentration, and the slope of the density gradient is determined by the time of centrifugation and the centrifugal field (see Fig. 4.2.2). Generally, the duration of separation procedures is selected so that organelles sediment to near equilibrium density (isopycnic conditions). Short runs (~30 min) enable large particles such as 1-µm-diameter secretion granules and mitochondria to approach equilibrium, whereas longer runs (up to 2 hr) may be required for smaller particles.

In the example presented in this protocol, Percoll gradients are used to purify secretion granules from exocrine glandular tissue and then to separate the granules into subpopulations according to the extent of their maturity (see Zastrow and Castle, 1987). The initial purification procedure is widely applicable to other types of granules.

Additional Materials *(also see Alternate Protocol 1)*

Percoll gradient solutions (see recipe)
Centrifuge and fixed-angle rotor (e.g., Beckman 70Ti or Type 50)
Centrifuge tubes (e.g., Beckman thick-walled polycarbonate bottle assemblies; 25-ml for 70Ti rotor or 10-ml for Type 50 rotor)
Refractometer

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

CU IAS - CENTRAL UNIVERSITY LIBRARY

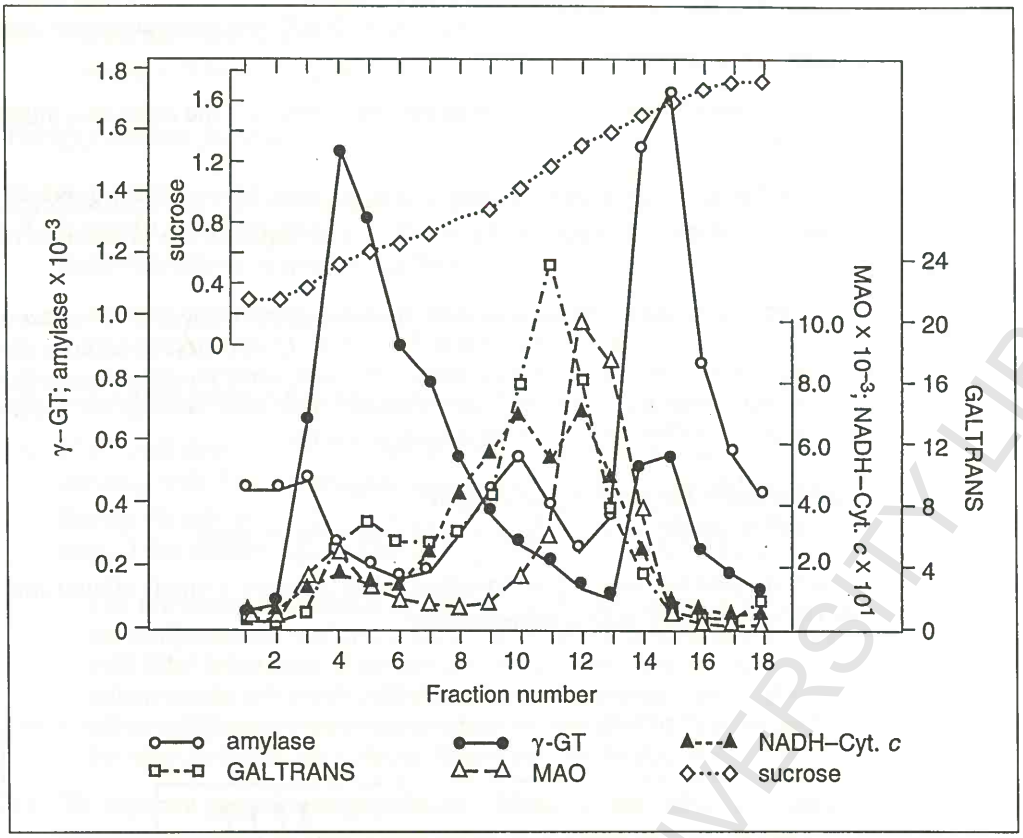


Figure 4.2.1 Distribution of marker activities for various organelles after rate zonal centrifugation of a postnuclear supernatant (PNS) from rat pancreas using the conditions specified in Basic Protocol 1. The density profile of the gradient after the run was determined by refractive index and plotted as sucrose (sucrose plus Ficoll) concentration; this is shown in the upper portion of the figure. The density profile of the gradient after the run was determined by refractive index and is shown in the upper portion of the figure. The activities that were assayed and the organelles that they mark are: amylase, a secretory protein concentrated in zymogen granules at the bottom of the gradient but also found in lower amounts in rough endoplasmic reticulum, Golgi, and at the top of the gradient (as a result of organelle lysis during homogenization); GALTRANS, galactosyltransferase, a marker for Golgi; γ-GT, gamma-glutamyl transferase, an integral component of plasma membrane but also of zymogen granule membranes (at lower concentration); MAO, monoamine oxidase, a marker of mitochondria (which is concentrated in outer mitochondrial membranes; also note the small amount near the top of the gradient resulting from organelle breakage); NADH–Cyt.*c*, NADH–cytochrome *c* reductase, a marker activity of rough ER and mitochondria. The distinct but overlapping distributions of these markers in this analytical centrifugation is apparent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Load and spin the samples

1. Remove parotid salivary gland from freshly sacrificed rats (0.4 to 0.5 g tissue per animal). Weigh gland, mince, homogenize, and filter through cheesecloth (see Basic Protocol 1, steps 1 to 3), except using homogenization medium containing DPPD (as in Alternate Protocol 1) and homogenizing at 12% (w/v).

   DPPD, an antioxidant, is added because it significantly reduces lysis of zymogen granules.

2. Centrifuge sample, rehomogenize (in homogenization medium with DPPD), and recentrifuge, then pool supernatants (see Basic Protocol 1, steps 4 and 5). Add MES

BOSTON UNIVERSITY LIBRARY

to 5 mM and EDTA to 1 mM. Filter and disperse with a Dounce homogenizer (see Basic Protocol 1, step 5).

3. Add the PNS to 2 vol of 60% Percoll/sucrose and mix. Measure the refractive index using a refractometer.

4. Load into tubes that are appropriate for the fixed-angle rotor that has been selected. Centrifuge the samples under appropriate conditions to separate the organelles of interest.

   To separate parotid gland secretion granules, appropriate centrifugation conditions are 30 min at 16,400 × g (e.g., 15,000 rpm in 70Ti), 4°C. Under these conditions the granules separate as a white band at the bottom of the tube, while the other organelles form a dense band near the top of the tube. A mostly clear zone between the bands corresponds to the shallow portion of the density gradient.

Collect bands and prepare for further purification

5. Collect the bands manually using a Pasteur or plastic transfer pipet.

6. For each fraction, mix by swirling or inverting, then remove a small aliquot and measure its refractive index with a refractometer.

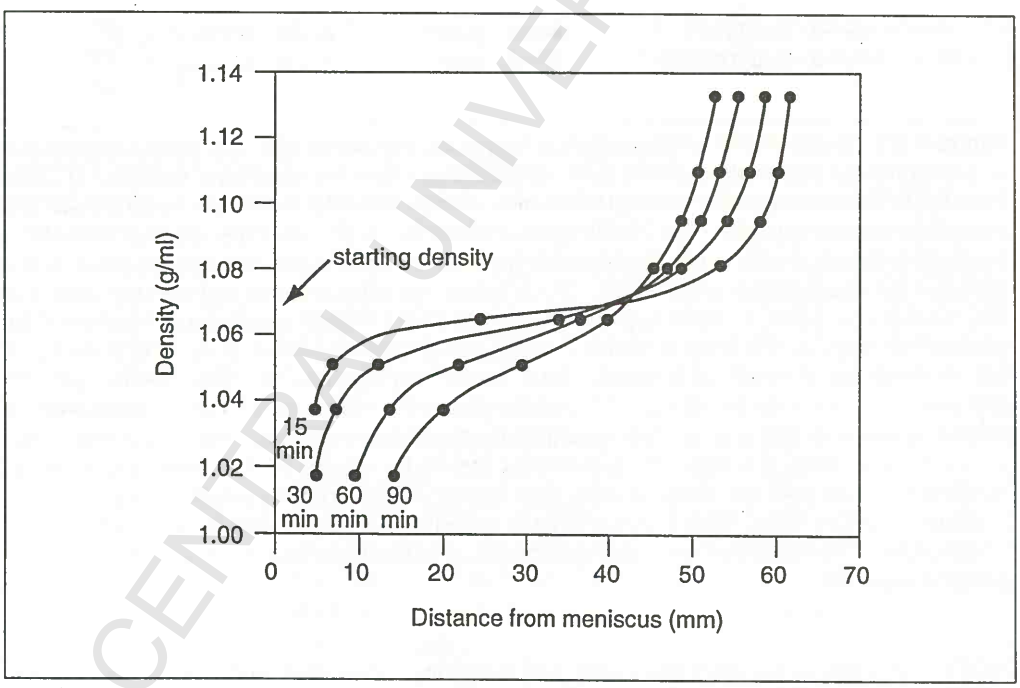


Figure 4.2.2 Percoll density gradient profiles as a function of time of centrifugation at 20,000 × *g* in a fixed-angle rotor. The starting density in the example shown is 1.07 g/ml. Note that with self-forming density gradients, the slope of the gradient begins as zero; during the run, the shape progressively changes as the Percoll sediments. Early in the run most of the tube contains a shallow gradient that is bordered by Percoll-depleted solution at the top and high-density pelleted Percoll at the bottom. As the run proceeds, the central resolving segment of the gradient becomes steeper and shorter. To produce an isoosmotic sucrose/Percoll solution of a desired density, use the formula:

$$\frac{V_x}{V_o} = \frac{\rho_o - \rho_x}{\rho_i - \rho_x}$$

where V_x = volume of 2.5 M sucrose, V_o = volume of Percoll, ρ_o = density of Percoll (1.13 g/ml), ρ_x = density of 2.5 M sucrose (1.315 g/ml), and ρ_i = desired density. Figure courtesy of Pharmacia Biotech.

CCU IASI CENTRAL UNIVERSITY LIBRARY

7. Adjust the refractive index back to the same value measured for the original diluted PNS by adding homogenization medium.

Purify fractions further

Example: For parotid-gland secretion granule fraction:

8. Load the adjusted granule fraction into centrifuge tubes and recentrifuge under the same conditions as used in the first run.

   This repeat centrifugation will float much of the residual contaminants away from the granule fraction.

9. Collect the granule fraction manually with a pipet.

10. If desired, process the purified granule fraction as a single organelle population by diluting with 3 vol homogenization medium, then pelleting the granules by centrifuging 45 min at 2000 × *g* (e.g., 5000 rpm in 70Ti), 4°C. Alternatively, proceed to step 11 for further separation of the granule fraction by Percoll centrifugation.

    The low-speed centrifugation removes most, but not all, of the Percoll and provides an easily resuspended pellet. Removal of Percoll is necessary because it can interfere with later processing of the sample. If necessary, further Percoll removal and purification can be achieved by dilution and recentrifugation under the same conditions, by ultracentrifugation on a step gradient as specified by Zastrow and Castle (1987), or by other methods (see Critical Parameters for further discussion).

11. To separate granule subpopulations: Measure the refractive index of the granule fraction collected in step 9 and adjust to a value of $n = 1.3586$ using either 0.3 M sucrose or 86% Percoll/sucrose, depending on whether refractive index must be adjusted downward or upward, respectively.

    The value of n will differ and must be determined empirically for granules from other sources.

12. Place the adjusted granule fraction in a new set of centrifuge tubes and centrifuge 30 min at 45,500 × *g* (e.g., 25,000 rpm in 70Ti or Type 50 rotor, depending on the volume), 4°C.

13. The granule fraction will now be subdivided into a minor upper band, a major lower band, and the clearer (but slightly turbid) zone between the bands. Collect the upper band, the intervening zone, and the lower band as separate fractions.

    Together these fractions comprise a continuum of increasingly dense granules (and increasing granule maturation), rather than discrete subpopulations. Evidently, the relative amounts of granules in each subfraction depend on the initial density (and thus refractive index) before the spin and on the centrifugation conditions (because Percoll gradients are self-forming and constantly change during the centrifugation).

15. Process the collected bands as in step 10 to remove Percoll. Proceed with analysis.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 3

GRADIENT FRACTIONATION USING D_2O

Deuterium oxide fractionation takes advantage of the fact that the density of D_2O is 1.1 g/ml—substantially closer than that of water to the density of biological organelles. The procedure is a form of equilibrium centrifugation in which D_2O (heavy water) is substituted for H_2O. This technique has not been used widely in subcellular fractionation, largely because deuterium oxide is quite expensive, and also because the practical value of the approach has fallen short of expectations. In some cases, however, it can be quite effective, so it should be considered as an option.

The procedure presented below (taken from Gumbiner and Kelly, 1981) illustrates use of a combined Ficoll/D_2O density gradient to purify and analyze endocrine secretion granules derived from a pituitary cell line.

Additional Materials (also see Alternate Protocol 1)

Mouse pituitary AtT-20/D16V cells
PBS (*APPENDIX 2E*) containing 4 mM EGTA (from 100 mM EGTA, pH 7.4), 37°C
Homogenization medium: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/10 mM HEPES (pH 7.4)/2 mM EGTA/1 mM EDTA, 4°C
Deuterium gradient stock solutions: 0.25 M ultrapure sucrose/10 mM HEPES (pH 7.4)/1 mM EDTA in D_2O (e.g., Aldrich), with and without 20% (w/v) Ficoll
8% (w/v) Ficoll 70 (Pharmacia Biotech) in homogenization medium

75-cm^2 tissue culture flasks
Dounce glass homogenizer with type B pestle (Kontes or equivalent)
Sorvall centrifuge and SS-34 rotor (or equivalent)
Gradient-forming device with capacity ≥15 ml/chamber (Hoefer Pharmacia or equivalent)
Peristaltic pump to form continuous gradient
Ultracentrifuge with swinging-bucket rotor (e.g., Beckman SW27 or SW28)
Automated gradient collection apparatus: Autodensiflow device (Buchler) or gradient displacement device (Nyegaard or MSE Scientific)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after harvesting cells should be conducted at 0°C (ice bucket) or 4°C (cold room).

Homogenize the cells and prepare the crude granule pellet

1. Grow cells almost to confluency in five 75-cm^2 tissue culture flasks (~5×10^6/flask). Harvest cells at 37°C in PBS/4 mM EGTA and pellet at 4°C.
2. Resuspend in 15 ml of 4°C homogenization medium and homogenize with six strokes in a Dounce homogenizer.

   A newer ball-bearing homogenizer designed by Balch and Rothman (1985) is now used widely for homogenizing cultured cells.

3. Pellet large debris and organelles by centrifuging 5 min at 10,000 × *g* (e.g., 9500 rpm in SS-34 rotor in a Sorvall centrifuge), 4°C.
4. Transfer the supernatant to a fresh tube and centrifuge 35 min at 30,000 × *g* (16,000 rpm), 4°C, to pellet the crude granule fraction.
5. Resuspend the crude granule pellet in 1 to 2 ml homogenization medium using a micropipettor until fine particulates are no longer seen. If necessary, homogenize

CU IASI CENTRAL UNIVERSITY LIBRARY

using a small-capacity Dounce homogenizer to achieve the desired uniform suspension. Maintain on ice while preparing the gradient.

Prepare the Ficoll-D_2O gradient

6. Using the gradient stock solutions made in D_2O (with and without 20% Ficoll), prepare layering solutions that are 17%, 14%, 11%, 9%, and 8% Ficoll.

7. Prepare a Ficoll step gradient (in 100% D_2O) at the bottom of each tube by layering the following solutions (from step 6) using either a micropipettor or plastic tissue culture pipet:

   1 ml 20% Ficoll
   2 ml 17% Ficoll
   2 ml 14% Ficoll
   2 ml 11% Ficoll
   1 ml 9% Ficoll.

   The volumes specified are for a 38-ml centrifuge tube, which is used in an SW27 or SW28 rotor.

8. Mix 8 ml of 20% Ficoll gradient stock solution with 12 ml homogenization medium to give 20 ml of 8% Ficoll in 40% D_2O medium.

9. Place 14.5 ml of the 8% Ficoll/40% D_2O solution in the back reservoir of a gradient-forming device. After filling the delivery tube between the chambers, place 14.5 ml of the 8% Ficoll in 100% D_2O gradient stock solution (from step 6) into the mixing chamber. Open the delivery tube, starting stirring the mixing chamber, and start the peristaltic pump to deliver a continuous gradient 40% to 100% D_2O in 8% Ficoll on top of the Ficoll step gradient.

Load the sample and spin

10. Load the resuspended crude pellet on the gradient and centrifuge 12 to 15 hr at 82,500 × *g* (e.g., 26,000 rpm in SW27 or SW28 rotor), 4°C.

    The centrifugation conditions listed (including tube size and rotor) are those reported by Gumbiner and Kelly (1981). It is very likely that this procedure can be successfully scaled down for use with smaller samples and smaller-volume gradients.

Collect fractions from the gradient and analyze

11. Collect 1.5-ml gradient fractions using one of the following methods:

    a. Collect top-first manually using a micropipettor.

    b. Collect top-first by an automated method: use either a Buchler Autodensiflow device or gradient displacement device (e.g., Nyegaard or MSE Scientific).

    c. Collect bottom-first by puncturing a hole in the bottom of the tube and collecting drops.

    The third method was used in the published procedure (Gumbiner and Kelly, 1981).

    Analyses for the secretion granule marker ACTH and total protein indicate that granules are highly purified, 50-fold on average, and are located within the lower Ficoll gradient/100% D_2O portion of the tube (centered at a density of 1.17 g/ml). Most other organelles are centered within the 8% Ficoll/D_2O gradient portion of the tube as judged by the concentration of protein in this region. Soluble protein remains at the top of the gradient.

    Lysosomes are a problematic contaminant of endocrine secretion granules (e.g., Roep et al., 1990; Loh et al., 1984) and have not been assayed specifically in this procedure.

BOSTON UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 4

GRADIENT FRACTIONATION USING NONIONIC IODINATED SOLUTES

Another strategy that is being used to achieve higher-density fractionation media with reduced osmolarity and viscosity involves supplementing or replacing sucrose with iodinated solutes. By far the most successful of these solutes are the nonionic derivatives of triiodobenzoic acid, metrizamide and Nycodenz (both available from Nyegaard). Metrizamide [2-(3-acetamido)-5-*N*-methylacetamido-2,4,6-triiodo(benzamido)-2-deoxy-D-glucose] and Nycodenz (*N,N'*-bis(2,3-dihydroxypropyl)-5-*N*-(2,3-dihydroxypropyl)acetamido-2,4,6-triiodoisophthalamide) are soluble in all aqueous media, stable over the pH range 2 to 12.5 and, unlike related ionic solutes, are not affected by low pH or the presence of divalent cations. These media are advantageous for fractionation under sterile conditions (Nycodenz can be autoclaved, whereas metrizamide must be sterile filtered). The differences in fractionation between these two media are negligible, and for most purposes, they are interchangeable. Note, however, that reported buoyant densities of selected organelles can differ slightly between the two (e.g., mitochondria has a density of 1.115 g/ml in metrizamide and 1.132 in Nycodenz). For a more extensive discussion of the properties of these solutes, see the monograph on centrifugation edited by Rickwood (1984).

This procedure describes purification of lysosomes from rat liver using a metrizamide gradient (Wattiaux et al., 1978); studies isolating other organelles by this method are noted in Background Information. In addition to preparing a high-quality fraction, these investigators show how results obtained with a continuous gradient are used to devise a step-gradient procedure for routine preparations.

Additional Materials (also see Alternate Protocol 1)

Rat starved overnight and freshly sacrificed
Homogenization medium: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals), 4°C
85.6% metrizamide (from 100% stock, Nyegaard)
Gradient stock solutions: for continuous gradients, 19.7% and 50% (w/v) metrizamide, pH 7.4; for step gradients, 19.78%, 24.53%, 26.34%, and 32.82% (w/v) metrizamide, pH 7.4 (adjust pH with 0.01 N NaOH)
Teflon-glass homogenizer (type C, 40 ml, A.H. Thomas)
Fixed-angle ultracentrifuge rotor appropriate to the size of the experiment (e.g., Beckman 70Ti)
Automated gradient collection apparatus: Autodensiflow device (Buchler) or gradient displacement device (Nyegaard or MSE Scientific)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare sample by homogenization and low-speed centrifugation

1. Remove liver from a freshly sacrificed rat that has been fasted overnight (to deplete glycogen stores). Weigh the liver, mince, and homogenize as for velocity centrifugation (see Basic Protocol 1, steps 1 to 2), except homogenize at ~20% (w/v) using a Teflon-glass homogenizer at 3000 rpm in 4°C homogenization medium (0.25 M sucrose).
2. Adjust the homogenate to 15% (w/v), then filter through cheesecloth. Centrifuge 15 min at 2,000 × *g*, 4°C.
3. Save the supernatant on ice and resuspend the pellet, adding homogenization medium

CU IAS / CENTRAL UNIVERSITY LIBRARY

to the original homogenate volume. Rehomogenize, then centrifuge again as in step 2.

4. Pool the two supernatants and centrifuge 30 min at 8,800 × *g* in a fixed-angle rotor (e.g., 10,000 rpm in 70Ti). Remove the supernatant, including the fluffy layer atop the pellet; resuspend the pellet in homogenization medium and repeat the centrifugation.

5. Resuspend the pellet in homogenization medium and mix with 2 vol of 85.6% metrizamide (to give a final 56.7% metrizamide load for the gradients).

Metrizamide and Nycodenz both absorb strongly in the UV range, complicating protein determinations in gradient fractions by OD_{280}. Also, they interfere with some protein assays (e.g., Lowry). To assay protein in various fractions, the samples can be precipitated with trichloroacetic acid and redissolved in 0.5 N NaOH. Metrizamide inhibits galactosyltransferase activity, so if this Golgi assay is to be performed, the fractions must be diluted, pelleted, and resuspended in a sucrose (or other noninhibitory) medium.

Prepare continuous metrizamide gradient and load sample

6. Using the gradient-forming device (Fig. 4.1.2), place the 50% metrizamide solution (density 1.280 g/ml) in the back reservoir and fill the delivery tube to the mixing chamber. Next place the 19.8% metrizamide solution (density 1.105 g/ml) in the mixing chamber. Set the outlet tube at the bottom of the centrifuge tube.

7. Open the delivery tube between the two chambers, start stirring in the mixing chamber, and pump the gradient (in low-density-first configuration).

Wattiaux et al. (1978) treats the continuous gradient as an analytical run and uses 5-ml gradients.

8. After the gradient is delivered, shut off the pump, taking care not to let air pass into the gradient from the outlet tube. Slowly withdraw the outlet tube from the centrifuge, taking care not to disturb the gradient.

9. Layer the sample in 56.7% metrizamide (final) beneath the gradient using either a Pasteur pipet or a syringe with long needle.

The reason for loading the gradient at the bottom is that mitochondria exhibit a higher equilibrium density when initially equilibrated in high-density metrizamide than when equilibrated in sucrose homogenization medium. This strategy is critical to achieving separation of mitochondria and lysosomes in the gradient.

10. Centrifuge the gradients 2½ hr at 108,000 × *g* (e.g., 28,000 rpm in SW28), 4°C.

Unload the gradients, measure refractive index, and assay marker activities

11. Working at 4°C, collect gradient fractions of equal volume using one of the following methods:

    a. Collect top-first manually using a micropipettor.
    b. Collect top-first by an automated method: use either a Buchler Autodensiflow device or gradient displacement device.
    c. Collect bottom-first by puncturing a hole in the bottom of the tube and collecting drops.

12. Measure the refractive index of each fraction. Using either a published table (available from Nyegaard) or standard curve constructed from a set of precisely made metrizamide stocks, construct a density profile for the gradient.

13. Using each gradient fraction plus an aliquot of the original gradient load, assay marker

BOYLAS / CENTRAL WINDSOR UNIVERSITY LIBRARY

activities for lysosomes (e.g., acid phosphatase, β-galactosidase), mitochondria (e.g., cytochrome oxidase), peroxisomes (e.g., catalase), and other organelles as desired.

14. From the activities present in each fraction, calculate the total recoveries of each of the enzymes present in the original gradient load. Plot the frequency of the components for each fraction as activity/(sum of the product of total enzyme activity × the density increment for each fraction).

Using these plots, it was found that peroxisomes equilibrate near the bottom of the gradient (median density 1.225 g/ml) with a portion of the catalase remaining in the load; mitochondria equilibrate near the middle (median density 1.162 g/ml); and lysosomes equilibrate in the upper region of the gradient (median density 1.133 g/ml).

Use step gradient to isolate specific organelles

Example: Isolate preparative levels of lysosomes

15. Based on the findings made using the continuous analytical metrizamide gradient (step 14), design an appropriate bottom-loaded step gradient. Resuspend the starting sample (light mitochondrial pellet) in homogenization medium, using 1 ml per 3 g original weight of liver, and mix with 2 vol of 85.6% (w/v) metrizamide (to give 56.7% load). Pipet 10 ml into a Beckman SW27 or SW28 tube, then overlay successively with 6 ml of 32.82% metrizamide (density 1.181 g/ml), 6 ml of 26.34% metrizamide (density 1.145 g/ml), 6 ml of 24.53% metrizamide (density 1.135 g/ml), and 9 ml of 19.78% metrizamide (density 1.109).

16. Centrifuge step gradient 2 hr at 95,000 × *g* (e.g., 27,000 rpm in SW27 or SW28), 4°C.

17. Collect bulk fractions from the top of the gradient using a pipet. Fraction 1 extends from the top to about two-thirds of the way through the 19.78% metrizamide layer; fraction 2 includes the 19.78%/24.53% interface (expected from step 14 to be enriched for lysosomes); fraction 3 includes the 24.53%/26.34% interface (expected from step 14 to be enriched for mitochondria); and fraction 4 includes the 32.82%/56.7% (load) interface (expected from step 14 to be enriched for peroxisomes).

18. Assay for marker activities and by electron microscopy to confirm that the fractions are enriched as expected.

Notably, Wattiaux et al. (1978) report that fraction 2 contained 10% to 12% of the lysosomal hydrolase activity of the total homogenate (a respectable yield) and had a specific activity for the lysosomal markers that on average was 74-fold higher than that of the homogenate (a very impressive purification).

BASIC PROTOCOL 3

GEL FILTRATION TO ISOLATE SECRETORY VESICLES

Gel filtration has not been applied widely in organelle purification; however, for small, homogeneous vesicles that can be included in the matrices of selected inert resins, it has proven to be valuable. As an example of this technique, the following protocol describes the use of Sephacryl S-1000 to fractionate the equivalent of a microsomal pellet obtained from a secretory vesicle–accumulating *sec6* mutant of *S. cerevisiae* (based on Walworth and Novick, 1987).

NOTE: During the first run of the column, substantial loss of material may occur, as there appears to be adsorption of lipid-rich vesicles to the column matrix. Adsorption can be substantially reduced by initially running sonicated phospholipid vesicles as a sample (this is also recommended when using controlled-pore glass). Alternatively, glyceryl CPG glass beads, which have a hydrophilic coating (polyethylene glycol or a similar polymer) that reduces surface adsorption, may be used.

CCU IASI CENTRAL UNIVERSITY LIBRARY

Materials

Saccharomyces cerevisiae sec6-mutant strain spheroplasted in 1.4 M sorbitol
Running buffer: 0.8 M sorbitol/10 mM triethanolamine/1 mM EDTA (pH 7.2; adjust pH with acetic acid)
1.5 × 90–cm to 1.5 × 100–cm column packed with Sephacryl S-1000 superfine resin (Pharmacia Biotech) and equilibrated with running buffer at 4°C

1-ml-capacity Dounce homogenizer
Centrifuges: low-speed and ultracentrifuge
Peristaltic pump capable of producing flow rates of ~10 ml/hr
Fraction collector

Additional reagents and equipment for gel-filtration chromatography (*UNIT 8.3*)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after harvesting cells should be conducted at 0°C (ice bucket) or 4°C (cold room).

Load and run the column

1. Lyse spheroplasted yeast in running buffer using a Dounce homogenizer.

   Conditions for spheroplasting can be found in Becker and Lundblad (1994). The running buffer is made with sorbitol instead of sucrose because yeast metabolize sucrose.

2. Centrifuge the lysate 10 min at 10,000 × *g*, 4°C, to remove mitochondria and larger particulates.
3. Remove the supernatant and centrifuge 1 hr at 100,000 × *g*, 4°C, to produce a microsomal pellet.
4. Resuspend the pellet in 600 µl running buffer and disperse using a Dounce homogenizer.
5. Remove running buffer from above the resin bed of the column and load the sample at 4°C.

   The column size suggested is suitable for most applications; a smaller-dimension column has been used by Barlowe et al. (1994) in fractionating small amounts of radiolabeled material.

   Additional information on column packing and setup may be found in UNIT 8.3.

6. Once the sample has entered the resin bed, overlay the bed with running buffer and, using a peristaltic pump, elute the column at a flow rate of ~10 ml/hr, collecting fractions continuously.

   80-drop (4-ml) fractions were collected by the investigators who designed this procedure.

Assay the column fractions and process those of interest

7. Vesicles small enough to enter the pores of the resin at least partially run in the included volume of the column; larger vesicles are excluded and appear in the void volume; and soluble proteins elute mostly after the vesicles. Locate the vesicle of interest in the column profile by assaying for a marker activity.

   A sample elution profile (Walworth and Novick, 1987) is shown in Figure 4.2.3.

   Invertase is used as a marker in the case of secretory vesicles (biosynthetically labeled secretory protein could also be used, as in Barlowe et al., 1994).

8. Pool the peak fractions and concentrate the vesicles for further study by centrifuging 1 hr at 100,000 × *g*, 4°C.

BCU IMS / CENTRAL UNIVERSITY LIBRARY

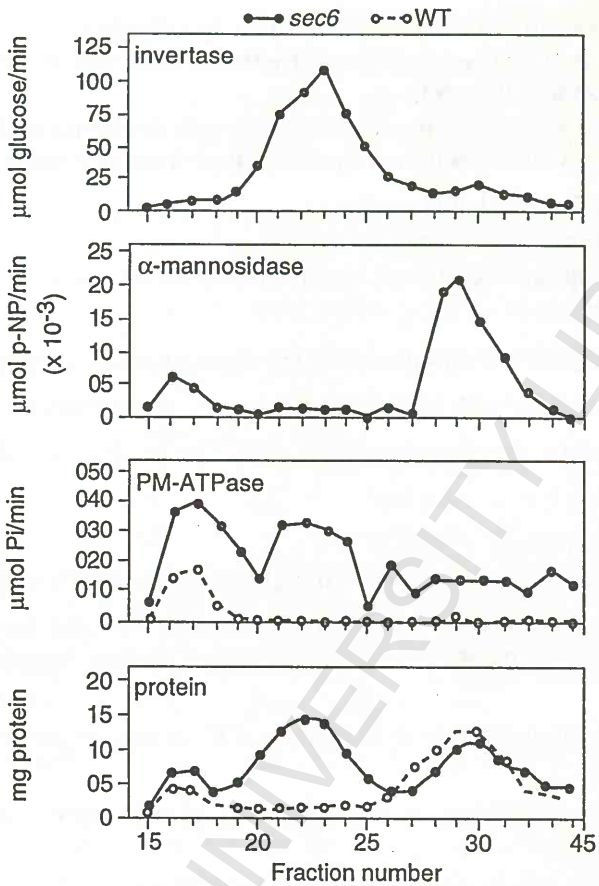


Figure 4.2.3 Distribution of organelle marker activities across eluted fractions from a Sephacryl S-1000 column on which a resuspended microsomal pellet from *sec 6-4* mutant (or wild-type) yeast was loaded. Fractions were assayed for invertase (a secretory protein), α-mannosidase (a vacuolar marker), plasma membrane ATPase, and protein. Invertase marks the distribution of secretory vesicles that accumulate in the *sec 6-4* mutant (but are absent in wild-type yeast). Part of the plasma membrane ATPase, but <1% of α-mannosidase, codistributes with invertase-containing secretory vesicles in the mutant. The protein indicates that significant protein elutes at the position of the vesicle peak for microsomes derived for mutant but not wild-type yeast. Reproduced from Walworth and Novick (1987) with permission of The Rockefeller University Press.

BASIC PROTOCOL 4

PREPARATIVE ELECTROPHORETIC SEPARATION

Electrophoretic purification can be used to separate complex mixtures of small vesicles that differ only modestly in size and density, because it takes advantage of the differences in their surface charges. This technique can be used preparatively or (as is more common) at the analytical level (Alternate Protocol 5), to separate and analyze organelles that have been selectively labeled. In this protocol, agarose electrophoresis using material derived from rat liver (Kedersha and Rome, 1986) is presented as an example of a preparative electrophoresis procedure. Although the protocol presented here used buffers of pH 6.5, the authors state that other pHs can be used, with resulting changes in the relative migration of different vesicles (as determined empirically).

Materials

Homogenization medium: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/50 mM 4-morpholinoethanesulfonic acid (MES), pH 6.5

ECU IAS/CENTRAL UNIVERSITY LIBRARY

6.25% sucrose/6.25% Ficoll 70/50 mM MES, pH 6.5
50 mM MES, pH 6.5
Running buffer: 50 mM MES (pH 6.5) with added sucrose as desired
0.15% to 0.2% Isogel agarose (FMC), or equivalent, in running buffer

Centrifuges: low-speed and ultracentrifuge
Standard flatbed agarose gel electrophoresis apparatus with 0.06-in. holes drilled in the cover to allow passage of inlet and outlet tubing from sample collecting wells (construct or purchase from Idea Scientific)
Standard 1.5-mm-thick plastic (siliconized) or Teflon gel comb
1.5-mm-thick plastic (siliconized) or Teflon eluting combs cut so that teeth are 3 mm wider than loading wells
0.062-in.-o.d. (1.57-mm) polyethylene tubing (Clay-Adams PE-160)
Peristaltic pump capable of flow rates of 0.2 to 15.0 ml/hr
Fraction collector

Additional reagents and equipment for preparing microsomal fractions (see Basic Protocol 1)

Prepare crude microsomal fraction

1. Prepare a crude rat liver microsomal fraction according to the procedure for differential centrifugation by velocity (see Basic Protocol 1), except using homogenization medium containing 50 mM MES. At the last step (see Basic Protocol 1, step 13), resuspend the crude microsome pellet in 6.25% sucrose/6.25% Ficoll 70/50 mM MES using a Dounce homogenizer.

2. Centrifuge the homogenate 40 min at 40,000 × *g*, 4°C, to pellet larger membrane vesicles.

3. Dilute the resulting supernatant with 4 vol of 50 mM MES or homogenization medium and centrifuge 2 hr at 100,000 × *g*, 4°C.

   Use homogenization medium (containing 0.25 M sucrose) in steps 3 and 4 if possible osmotic damage is a concern.

4. Resuspend the pellet (concentrated microsomes) in ≤0.8 ml of 50 mM MES (pH 6.5) or homogenization medium, and microcentrifuge 3 min at top speed to remove aggregates. Save the sample on ice until it can be applied to the gel (step 9).

   Samples for preparative electrophoresis can contain as much as 40 to 80 mg/ml protein.

Prepare and load the gel

5. Melt 0.2% Isogel agarose in running buffer using a microwave oven.

   FMC Isogel agarose is preferred because it has a low content of charged groups; it is usually used at concentrations of 0.15% to 0.2% but may be applied to separation of larger organelles (>200 nm diameter) when used at lower concentrations (<0.15%).

6. Cast the gel, with both the sample loading and eluting combs carefully prepositioned, and allow to solidify ≥2 hr at 4°C.

   Eluting combs are used to facilitate recovery of any particles that have diffused laterally (perpendicular to the field) during electrophoresis. Siliconization of plastic combs is recommended to minimize tearing of low-percentage agarose gels.

7. Remove eluting comb and immediately fill the eluting well(s) with running buffer. Keep the sample comb in place until just before loading.

BOUVIER CENTRAL UNIVERSITY LIBRARY

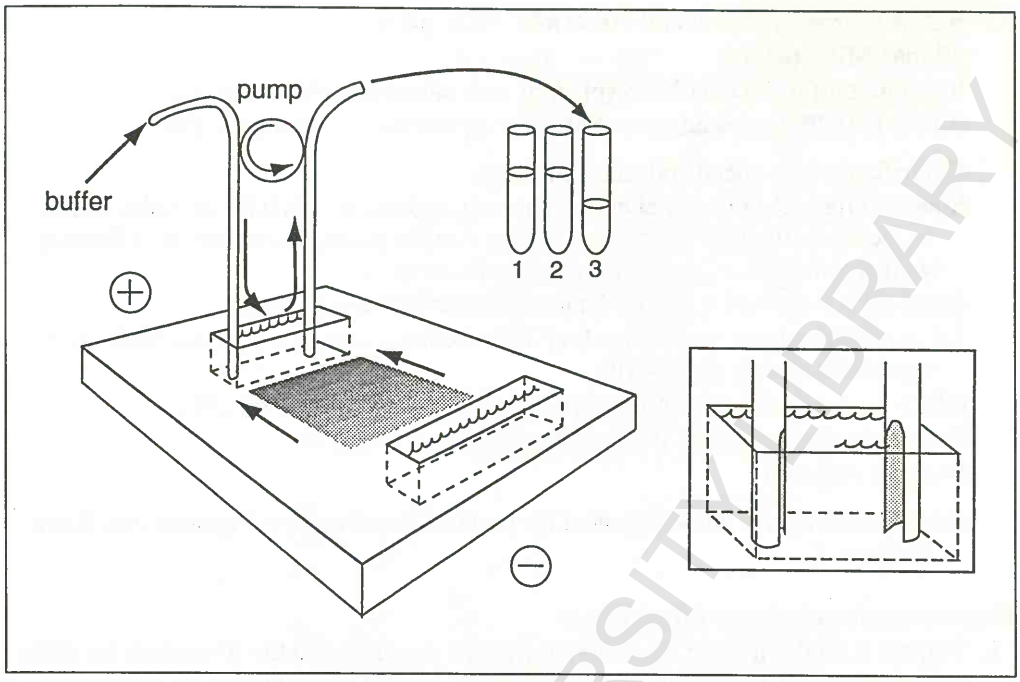


Figure 4.2.4 Schematic showing the horizontal agarose gel system devised by Rome and colleagues. Note the loading well, near (–) electrode, and the eluting well with continuous flow, near (+) electrode. The inset at the lower left shows the configuration of inlet and outlet tubing in the eluting well; note the cutouts from the walls of the tubing within the well. Figure courtesy of L. Rome with permission of Analytical Biochemistry.

8. Place the gel in the apparatus, taking care to align the elution wells with the holes in the lid.

   Electrophoresis is performed at 4°C, so all further operations should be conducted at this temperature (in a cold box or cold room).

9. Remove sample comb and load sample (from step 4).

10. Fill buffer reservoirs with running buffer until just level with the upper surface of the gel.

11. Attach the lid of the apparatus. Connect two pieces of 0.062-in.-o.d. polyethylene tubing whose ends have been cut away on one side to a peristaltic pump. Insert tubing through the holes in the lid and fully into the eluting well (see Fig. 4.2.4).

    The tubing provides buffer inlet and sample collection from the eluting well(s). The cutaway sides on the ends of the tubing should face one another for efficient buffer delivery and sample collection.

Run the gel and continuously collect the eluted fractions

12. Connect the electrodes, with the direction of electrophoresis running toward the (+) electrode (see Fig. 4.2.4), and turn on the power supply and the peristaltic pump. Set the current at 35 mA (~40 V). Adjust the flow rate on the peristaltic pump empirically.

    For example, for a 2-cm-wide elution well set 4 cm downfield from the sample well in a 1-cm-thick gel and an electric field of 1 V/cm, a flow rate of 4 ml/hr is reported to give quantitative elution.

13. Place a flask of warm water on the apparatus during the run to prevent condensation on the underside of the lid.

CU IAS / CENTRAL UNIVERSITY LIBRARY

14. Carry out the electrophoresis for 16 hr.

15. Collect 3 to 4 ml fractions (generally, 30 to 45 total fractions) and assay initially by analytical SDS-PAGE (*UNIT 10.1*) or by more specific assays for selected markers (e.g., dot blotting with an antibody).

    The duration of electrophoresis and the size of fractions collected will generally have to be adjusted empirically to suit the particular fractionation being performed.

15. Concentrate pooled fractions by centrifuging 30 min at 100,000 × *g*, 4°C. Resuspend the pellets in an appropriate buffer for further analysis.

ELECTROPHORETIC SEPARATION ON A DENSITY GRADIENT

ALTERNATE PROTOCOL 5

As indicated in the initial discussion of electrophoretic separations above (see Basic Protocol 4), electrophoresis has proven to be very useful for separating specific membrane populations from among similar-sized vesicles where there are unique surface charge characteristics. In this protocol, the unusually high negative surface charge on endosomes has enabled the identification of a novel population that accumulates MHC class II molecules en route to the cell surface. Its distribution following electrophoresis differs from the distributions of markers of other organelles but corresponds to the distributions of invariant chain processing and antigenic peptide loading (Tulp et al., 1994).

Additional Materials (also see Basic Protocol 4)

Rat starved overnight and freshly sacrificed
Electrophoresis buffer: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/10 mM triethanolamine/10 mM acetic acid/1 mM EDTA, pH 7.4
Trypsin, TPCK-treated (Worthington Biochemical or equivalent)
Soybean trypsin inhibitor
10%, 6%, 5%, and 4% (w/v) Ficoll 70, prepared in electrophoresis buffer

Electrophoresis apparatus: cylindrical electrophoresis chamber with dimensions of 5 cm high and 2.2 cm i.d. with circular palladium and platinum electrodes and cation-permeable membrane (Ionics, Inc.; see Tulp et al., 1993) *or* Polyprep 2000 apparatus (Buchler)
Gradient-forming device capable of forming linear concentration gradients of 10 ml volume (Hoefer Pharmacia)
1-mm-gauge stainless steel mesh

Additional reagents and equipment for preparing microsomal fractions (see Basic Protocol 1)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue, including forming and running the gradient, should be conducted at 4°C.

Prepare crude microsomal fraction

1. Remove liver from freshly sacrificed rat that has been fasted overnight (to deplete hepatic glycogen). Weigh, mince, homogenize, filter, and centrifuge at low speed as for centrifugation by differential velocity (see Basic Protocol 1, steps 1 to 4), using electrophoresis buffer as the homogenization medium and centrifuging 15 min at 840 × *g* to remove nuclei and larger debris.

2. Recover the supernatant (PNS), assay its protein content, and add trypsin (25 µg/mg PNS protein). Incubate 5 min at 37°C.

   Trypsin treatment is employed by most investigators who use electrophoresis to

BROWN UNIVERSITY LIBRARY

fractionate cellular organelles by electrophoresis. Treatment with trypsin enables compartments of the endocytic pathway to be resolved from other microsomal constituents (Marsh et al., 1987).

3. Stop the digestion by adding soybean trypsin inhibitor (100 µg/mg protein) and cooling on ice. Centrifuge 60 min at 100,000 × *g*, 4°C.

4. Resuspend pellet (crude microsomal fraction) in 0.5 ml of 5% Ficoll 70. Maintain on ice until it can be loaded for electrophoresis (step 7).

Set up the electrophoresis apparatus and load with the gradients and sample

5. Fill electrophoresis chamber with 10% Ficoll 70 via the inlet at the bottom. Fill the lower buffer reservoir with electrophoresis buffer.

A schematic diagram of an apparatus for electrophoretic separation is shown in Figure 4.2.5.

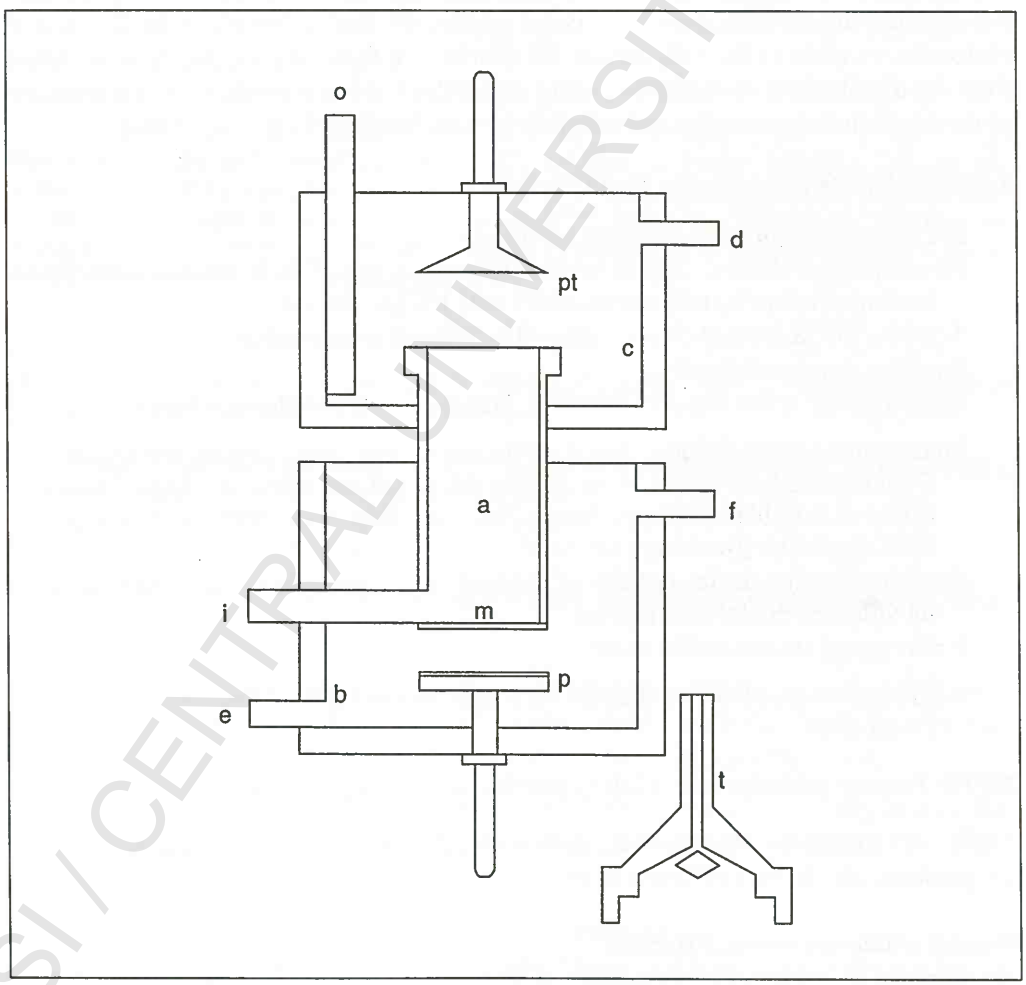


Figure 4.2.5 Apparatus used for electrophoresis on a density gradient. **a**, separation column (5 cm tall; 2.2 cm wide) containing the density gradient; **b**, bottom reservoir; **c**, top reservoir; **d**, outlet for the top reservoir; **e**, inlet for the bottom reservoir; **f**, outlet for the bottom reservoir; **i**, inlet for the separation column; **m**, cation-permeable membrane (part 61 AZL-386, Ionics Inc., MA, USA), which is held in place by a threaded fixture; **o**, inlet for the upper reservoir; **p**, circular (3.8-cm^2) palladium anode; **pt**, platinum cathode (3.8 cm^2); **t**, top cone with flow deflector (the inlet through which the Ficoll gradients are delivered). After removing **t**, the 1-mm-mesh stainless steel screen is placed atop **a** while the upper reservoir is filled, and the removed. After removing the screen, **pt** is installed and the electrophoretic separation is started. Reproduced from Tulp et al., 1993, with permission of VCH Verlagsgesellschaft.

ICU IASI / CENTRAL UNIVERSITY LIBRARY

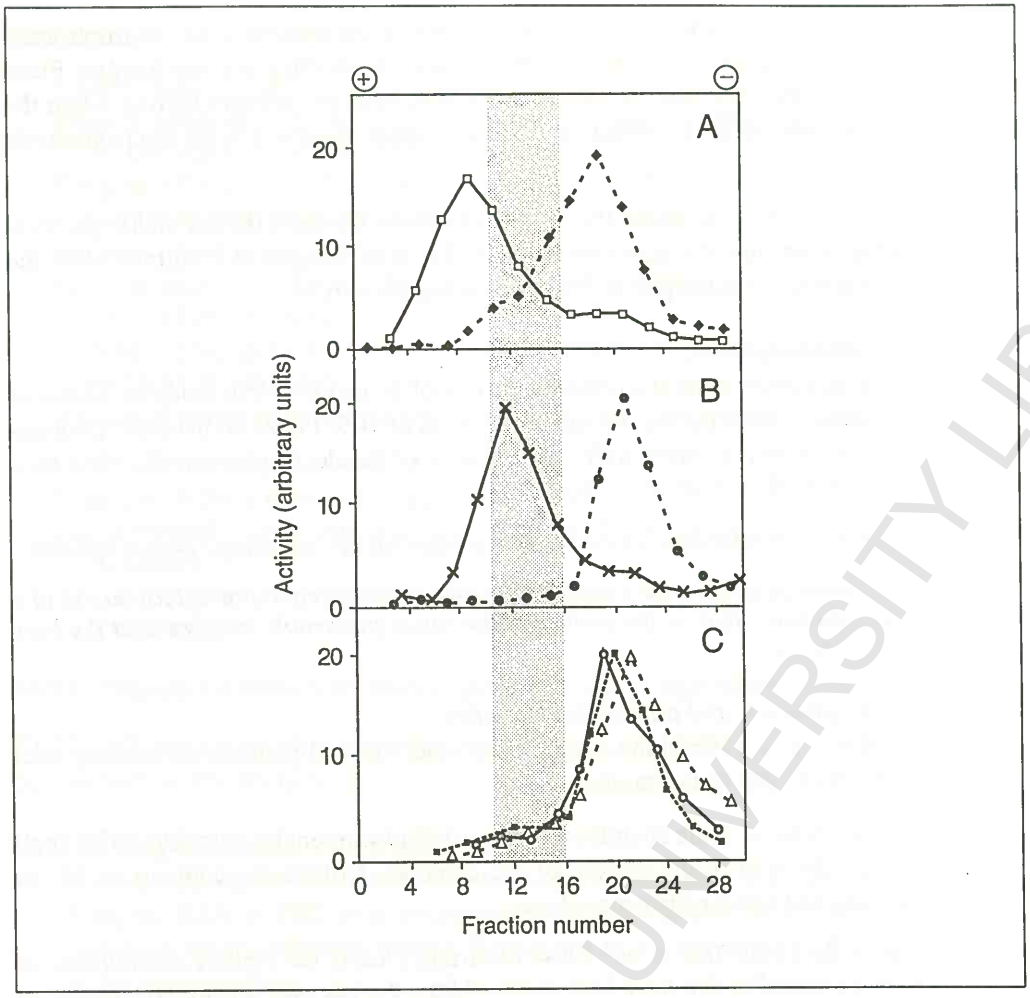


Figure 4.2.6 Results from density gradient electrophoresis of various organelles derived from cultured melanoma cells. (**A**) Distribution of horseradish peroxidase (HRPO) after endocytosis at 37°C. Filled diamonds show distribution of early endosomes (4 min HRPO uptake) and open squares show distribution of late endosomes (4 min HRPO uptake followed by 30 min chase at 37°C). (**B**) Distribution of the lysosomal hydrolase β-hexosaminidase (×) and plasma membranes that were marked in intact cells by surface iodination at 4°C using lactoperoxidase (filled circles). (**C**) Distributions of Golgi membranes (galactosyltransferase activity; open circles); ER (biosynthetically labeled by a 2-min pulse with ^{35}S methionine/cysteine; filled squares), and total protein (open triangles). The shaded region indicates where the bulk of a novel compartment containing MHC class II molecules are concentrated. Data are from Tulp et al. (1994); reprinted with permission from *Nature* (copyright 1994 Macmillan Magazines Ltd.).

6. Attach the gradient maker to the top inlet of the chamber and introduce a gradient (total volume 6 ml) descending from 10% to 6% Ficoll 70 above the 10% Ficoll. The gradient is formed and introduced into the electrophoresis chamber by displacement of the 10% Ficoll solution, which is allowed to drip slowly out of the lower inlet.

7. Load the crude microsomal sample (from step 4) above the Ficoll gradient, again by displacement of the 10% Ficoll solution, which is allowed to drip out of the lower inlet of the electrophoresis chamber.

8. Above the loaded sample, form and load a 9.5-ml continuous gradient of 4% to 0% Ficoll 70 by displacement of the 10% Ficoll at the bottom of the chamber. Upon completion, remove the loading inlet from above the electrophoresis chamber.

BODLEIAN LIBRARY UNIVERSITY

9. Allow 10% Ficoll 70 solution to drip from the inlet at the bottom of the electrophoresis chamber until the meniscus is level with the top of the electrophoresis chamber. Place 1-mm-gauge stainless steel screen over the meniscus (to prevent mixing when the upper reservoir buffer is added) and fill the upper reservoir with electrophoresis buffer.

10. Remove the screen and attach the top (+) electrode. Connect the peristaltic pump to the buffer reservoirs; the upper reservoir buffer is exchanged at 1 ml/min while the lower reservoir is exchanged at 5 ml/min during the run.

Start the electrophoresis

11. Carry out electrophoresis at a constant current of 10 mA (3 V/cm field) for a total of 40 min. At two times during the run, allow 6 ml of 10% Ficoll 70 in electrophoresis buffer to drip slowly (1 ml/min) from the bottom of the electrophoresis chamber, then replace with 6 ml of fresh 10% Ficoll 70.

Note that electrophoresis of vesicles with a negative surface charge occurs upward.

Replacement of some of the Ficoll 70 is a precaution taken to prevent formation of a low-conductivity layer at the surface of the cation-permeable membrane at the base of the chamber.

Stop the electrophoresis and unload the chamber

12. After ending the run, drain the upper buffer reservoir and reattach the loading inlet above the electrophoresis chamber.

13. Unload the contents of the chamber by upward displacement by pumping 10% Ficoll 70 through the inlet at the bottom of the chamber. Collect the contents in 25- to 30-fractions and use for further analysis.

Figure 4.2.6 (from Tulp et al., 1994) illustrates clearly the distinct distributions of different organelles that have been resolved from a crude microsomal preparation of melanoma cells by density gradient electrophoresis.

BASIC PROTOCOL 5

SOLID-PHASE IMMUNOADSORPTION

Solid-phase immunoadsorption using specific antibodies is a powerful technique for isolating cellular organelles that can be used on both the preparative and analytical levels. The procedure requires a high-quality antibody that reacts to an antigenic epitope exposed on the surface of the organelle of interest; the antigen must be reasonably specific to the organelle of interest. The antibody is adsorbed to a support (e.g., fixed *Staphylococcus aureus* cells, protein A–Sepharose, or magnetic beads), and the organelles of interest bind to the antibody support.

The procedure presented below describes use of antibody-coated magnetic beads, an approach developed for organelles by Gruenberg and Howell (1986). An antibody to the cytoplasmic tail of the polymeric immunoglobulin receptor is bound to the beads coated with secondary antibody; it is then used to isolate relatively large organelles—stacked elements of the Golgi complex—and vesicles that can be induced to bud in vitro from the stacked Golgi elements (Salamero et al., 1990).

CU IASI CENTRAL UNIVERSITY LIBRARY

Materials

5 mg/ml BSA in PBS
Primary antibody that selectively recognizes an antigenic epitope present on the exposed surface of the vesicle of interest
5 mg/ml BSA in 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/0.1 M potassium phosphate (pH 6.65)
Rat starved overnight and freshly sacrificed
Homogenization medium: 0.5 M sucrose/0.1 M potassium phosphate (pH 6.65)/5 mM $MgCl_2$
1.25 M, 1.3 M, and 1.4 M sucrose (ultrapure) in 0.1 M potassium phosphate (pH 6.65)/5 mM $MgCl_2$
4.5-μm-diameter magnetic beads coated with sheep anti-mouse Fc antibody (Dynal)
Magnetic collection device (e.g., Dynal)
End-over-end rotator capable of slow rotation (2 to 4 rpm; e.g., Pelco rotator, modified)
10-ml and 38-ml centrifuge tubes suitable, respectively, for Beckman Type 50 and SW28 rotors

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Coat the magnetic beads with primary antibody

1. Wash an appropriate amount of 4.5-μm-diameter magnetic beads twice with 1 ml of 5 mg/ml BSA in PBS in a microcentrifuge tube. Collect beads with a magnetic collection device between washes. Resuspend to a concentration of 2 mg beads/ml in the same medium.

   Coating the beads with antibody (steps 1 to 4) is usually done the day before the rest of the experiment. The beads are treated with protein to block nonspecific adsorption. Most procedures call for using bovine serum albumin (5 mg/ml in PBS or in buffered 0.25 M sucrose); one recent protocol recommends use of fetal bovine serum (3% in PBS) as the best way to reduce nonspecific binding (Saucan and Palade, 1994).

   Magnetic concentration is very quick and requires <1 min per wash. Dynal markets a number of magnetic collection devices, one of which is appropriate for microcentrifuge tubes. Alternatively, a plate magnet (Edmund Scientific) can be used beneath a cell culture dish (Howell et al., 1989).

   In addition to 4.5-μm beads, Dynal also markets 2.8-μm beads, beads coated with anti-mouse antibody, and tosyl-activated beads that can be custom-coupled to the antibody of choice by the investigator using a very simple procedure described by the manufacturer.

2. Add primary antibody (in 1 ml of 5 mg/ml BSA in PBS) to 2-fold excess over the amount of bound secondary antibody. Incubate 4 hr to overnight at 4°C in a microcentrifuge tube on an end-over-end rotator set at 2 to 4 rpm.

   Information regarding the amount of secondary antibody bound to the magnetic beads is provided by the manufacturer.

   A rotator for electron microscopy samples available from Pelco can be easily modified for the slow rotation needed for this protocol.

3. Wash the beads by adding 1 ml of 5 mg/ml BSA in PBS, 4°C, then collecting the

BOSTON UNIVERSITY LIBRARY

beads magnetically or by microcentrifugation (30 sec at 10,000 rpm). Repeat twice with 5 mg/ml BSA in 0.25 M sucrose/0.1 M potassium phosphate, pH 6.65.

In general, a control for nonspecific binding can be run in parallel using beads that have been coated with an irrelevant primary antibody.

Prepare sample for immunoadsorption

Example: Prepare sample enriched in stacked Golgi complexes from rat livers:

4. On the same day the immunoadsorption experiment is to be done, sacrifice a rat that has been starved overnight (to deplete hepatic glycogen) and remove liver. Weigh liver, mince, homogenize, and filter as for velocity differential centrifugation (see Basic Protocol 1), except homogenize at 20% (w/v) in 0.5 M sucrose/0.1 M potassium phosphate/5 mM $MgCl_2$ homogenization medium.

   This method for preparing a partially purified fraction enriched in stacked Golgi complexes from rat livers is based on the procedure of Leelavathi et al. (1970).

5. Prepare a postnuclear supernatant (PNS) by centrifuging sample as for differential centrifugation by velocity (see Basic Protocol 1, step 4).

6. Place a 4-ml cushion of 1.3 M sucrose/0.1 M potassium phosphate/5 mM $MgCl_2$ into a 10-ml centrifuge tube, and layer 6 ml PNS on top. Centrifuge 60 min at 105,000 × *g* (in Type 50 rotor), 4°C.

7. Collect the band at the surface of the 1.3 M sucrose layer and adjust to 1.1 M sucrose using homogenization medium.

8. Prepare a second set of discontinuous gradients containing 6 ml each of 1.4 M, 1.3 M, and 1.25 M sucrose in 0.5 M potassium phosphate/5 mM $MgCl_2$. Layer 6 ml of the adjusted sample (step 7) over each gradient, overlay with homogenization medium to fill the tubes, and centrifuge 90 min at 80,000 × *g* (in SW28 rotor), 4°C.

9. Collect the Golgi fraction at the interface between the load and the 1.1 M sucrose layer. Adjust to 0.25 M sucrose and 5 mg/ml BSA. Place on ice.

   The Golgi fraction is now ready to be used for immunoadsorption (which should be done on the same day). The fraction should be assayed for protein content and maintained on ice until use.

Adsorb organelles to antibody-coated beads

10. Add a sample of the organelle fraction (~0.5 mg protein in 0.3 to 0.4 ml of 0.25 M sucrose/5 mg/ml BSA) to 1 mg antibody-coated magnetic beads in 0.6 to 0.8 ml of 5 mg/ml BSA in PBS. Incubate 3 hr at 4°C with end-over-end rotation to allow binding to occur.

    Generally, the optimal ratio of input fraction to coated beads is determined in a preliminary binding experiment in which several 1-mg aliquots of beads are incubated with increasing amounts of the organelle fraction to determine the saturation point of binding. Bound and free organelles can be assessed by immunoblotting. For functional studies involving vesicle budding from adsorbed organelles, it is necessary to saturate antigen binding sites so that the vesicles can't bind once they have budded. Salamero et al. (1990) reached saturation at 0.5 mg (protein) of Golgi fraction per mg of beads.

Wash away unadsorbed organelles

11. Wash the coated beads with 0.25 M sucrose/5 mg/ml BSA. Repeat as many times as needed to remove all unadsorbed organelle, collecting the beads magnetically between washes.

    The number of washes needed should be determined in the preliminary immunoblotting experiment mentioned above.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

For adsorbed small vesicles (e.g., endosomes), the concentration device built for microcentrifuge tubes is satisfactory. For large organelles such as the stacked Golgi complexes, it is preferrable to use culture dishes and a plate magnet so that the adsorbed organelles are not damaged or dislodged by mechanical shear.

12. Use the immunoadsorbed fraction for further studies.

Examples of further applications include treatment with SDS solubilization buffer to desorb bound protein for SDS-PAGE and immunochemical analysis; suspension in fixative followed by preparation for electron microscopic analysis; incubation with a second primary antibody followed by a colloidal gold–conjugated secondary antibody to test for co-localization with the immunoadsorbing antigen by immunoelectron microscopy (e.g., Laurie et al., 1993); and functional studies of vesicle trafficking (Salamero et al., 1990).

PURIFICATION BY LECTIN ADSORPTION

BASIC PROTOCOL 6

Lectin adsorption has proven to be particularly useful for plasma membrane isolation and to a lesser degree for isolation of endosomes and lysosomes that have loaded by endocytosis of lectin conjugates at the cell surface. The latter procedures fall into the category of fractionation by density perturbation (Basic Protocol 7). In some procedures, the lectin conjugate is added to intact cells, thereby guaranteeing specific binding to plasma membrane; this approach appears to be applicable to cultured cells but not to tissues. In other procedures that are applicable to tissue samples, immobilized lectins can be used to adsorb selectively to plasma membrane vesicles in cell homogenates or partially purified fractions (e.g., crude microsomes). This strategy is possible because following cell or tissue disruption during homogenization, plasma membranes vesiculate with the ectodomain (glycosylated surface) facing outward, whereas most cytoplasmic organelles retain a cytosolic-surface-outward orientation. Once the fraction is bound to the lectin and washed free of unbound organelles, it can be eluted by competing the interaction between lectin and vesicles using the appropriate hapten sugar.

In this protocol, wheat germ agglutinin (WGA) affinity chromatography is used to purify apical plasma membranes from homogenates of pancreatic acinar cells (Paul et al., 1992). WGA is a lectin that binds both *N*-acetylglucosamine (GlcNAC) and *N*-acetylneuraminic acid (Neu5Ac) residues in glycoproteins, and interactions with WGA can be competed with free *N*-acetylglucosamine, which is relatively inexpensive. Additional guidelines for lectin affinity chromatography are presented in *UNIT 9.1*.

Materials

Freshly sacrificed Sprague-Dawley rat (200 to 275 g), fasted overnight
Homogenization medium (see recipe; add 0.1 mM PMSF just before fractionation), 0°C
0.2 M KCl/0.2 M NaBr/50 mM Tris·Cl, pH 6.8
50 mM Tris·Cl, pH 6.8
Wheat germ agglutinin (WGA) conjugated to Sepharose (Pharmacia Biotech), preequilibrated with 50 mM Tris·Cl, pH 6.8
0.1 M *N*-acetylglucosamine (e.g., Sigma) in 50 mM Tris·Cl, pH 6.8

Motor-driven Teflon-glass homogenizer
Cheesecloth
Centrifuge capable of producing a field of 25,000 × *g* and equipped with a fixed-angle rotor
Centrifuge tubes
5-ml chromatography column

BOSTON CENTER UNIVERSITY LIBRARY

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare a crude plasma membrane fraction

1. Remove pancreas from rat, weigh (generally 0.6 to 1.0 g tissue per rat), submerge in 0°C homogenization medium, and mince coarsely with scissors.

   Paul et al. (1992) used 6 g of pancreas. More than one rat may be required.

2. Homogenize at 10% (w/v) in homogenization medium by 10 strokes with a Teflon-glass homogenizer. Filter the homogenate through four layers of cheesecloth. Remove a small aliquot (~1% of total volume) for assay of marker activities.

3. Centrifuge 25 min at 120 × *g*, 4°C, to remove large cellular debris (most nuclei and larger).

4. Collect the supernatant and centrifuge 15 min at 13,000 × *g*, 4°C, to pellet zymogen granules and mitochondria.

5. Collect the supernatant and centrifuge 60 min at 25,000 × *g*, 4°C, to pellet plasma membranes and other vesicles.

6. Discard the supernatant and resuspend the pellet in 1 to 2 ml of 0°C homogenization medium. Remove an aliquot (~1% of total volume) and assay for marker activities.

   Marker assays include: (1) γ-glutamyl transferase, an apical plasma membrane marker (Orlowsky and Meister, 1963); (2) Na,K-ATPase, a basolateral plasma membrane marker (Baykov et al., 1988); and (3) alkaline phosphatase, a duct cell plasma membrane marker (Emmelot and Bos, 1966).

   It is reported that 65% of the marker activity of the apical plasma membrane, γ-glutamyl transferase, is present in this fraction (Paul et al., 1992).

   If activities are being followed during the course of purification, assaying all fractions generated at each step is recommended so that it will possible to account for all activity.

7. Dilute the rest of the resuspended membranes 10-fold with 0.2 M KCl/0.2 M NaBr/50 mM Tris·Cl, pH 6.8. Resediment by centrifuging 60 min at 25,000 × *g*, 4°C.

   NaBr is a chaotropic agent that has been used frequently in fractionating pancreas to promote desorption of extrinsic proteins from membrane vesicles.

8. Resuspend the pellets in 3 ml of 50 mM Tris·Cl, pH 6.8. Mix with 3 ml of WGA-Sepharose preequilibrated in 50 mM Tris·Cl, pH 6.8. Gently mix for 30 min in a closed tube, then transfer the entire mixture to a 5-ml chromatography column.

9. Wash the column after loading with at least 15 ml (5 vol) of 50 mM Tris·Cl, pH 6.8.

10. Elute the bound vesicles using 0.1 M *N*-acetylglucosamine in 50 mM Tris·Cl, pH 6.8. Concentrate the membrane-containing fractions as necessary by centrifuging 60 min at 25,000 × *g*, 4°C.

    Most of the membrane vesicles elute within 2 to 4 ml. This can be confirmed by assaying γ-glutamyl transferase of the fractions (a very easy spectrophotometric assay; Orlowski and Meister, 1963).

    Measurements of the specific activity of γ-glutamyl transferase (per protein) in the homogenate and the final purified membranes indicates that the purification is 75-fold, which is excellent for plasma membranes.

SCU IAS / CENTRAL UNIVERSITY LIBRARY

DENSITY SHIFT USING DIGITONIN

Density-shift techniques have found use in both analytical and preparative centrifugation procedures. In this approach, the sample to be separated is subjected to a treatment that differentially alters the densities of the component organelles, which are then separated by centrifugation. The following density-shift technique (based on the work of Amar-Costasec et al., 1974) involves the use of digitonin and capitalizes on differences in cholesterol content in the membranes of various cell organelles. Controlled, sublytic concentrations of digitonin are applied to the membranes, leading to the formation of 1:1 molecular complexes with cholesterol, and these complexes result in increased buoyant density during cell fractionation. The magnitude of the effect varies depending on the cholesterol content of different organelle membranes; consequently, different organelles exhibit distinct density shifts as judged from the distribution of marker activities. The example given is for microsomes from rat liver but can be adapted to membranes from other sources.

Materials

Rat starved overnight to deplete glycogen stores and freshly sacrificed
Homogenization medium: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/3 mM imidazole (pH 7.4)
20 mg/ml digitonin stock solution (see recipe)
10% (w/v) Triton X-100
Digitonin density shift gradient stock solutions (see recipe)
Sample for treatment and analysis

Motor-driven Teflon-glass homogenizer (type C, A.H. Thomas)
Cheesecloth
Low-speed refrigerated centrifuge
Ultracentrifuge with fixed-angle (e.g., Beckman 70Ti), and swinging-bucket rotors (e.g., Beckman SW28)
Ultracentrifuge tubes to fit rotors
Gradient-forming device (e.g., Hoefer Pharmacia)

Additional reagents and equipment for microsomal fraction preparation (see Basic Protocol 1)

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare crude microsomal sample

1. Remove liver from a sacrificed rat (6 to 7 g tissue) and prepare total microsomal preparation as in differential centrifugation by velocity (see Basic Protocol 1, steps 1 to 8 and 11 to 12), except using homogenization medium containing 3 mM imidazole.

   Imidazole (pH 7.4) was the buffer used when the procedure was devised; MOPS or HEPES can be substituted at the same concentration.

2. Resuspend the microsomal pellet thoroughly in homogenization medium using a Teflon-glass homogenizer and wash by repelleting under the same conditions used to pellet the microsomes initially (see Basic Protocol 1, steps 11 and 12).

3. Resuspend the pellet (total washed microsomes) in homogenization medium, using 1 ml medium per 1 g liver used in step 1 and keep on ice until use.

BCM INST CENTRAL UNIVERSITY LIBRARY

Determine the digitonin concentration to use for density shift by dose-response assay

4. For liver microsomes, proceed to steps 6 to 8; for membranes from other tissues, determine the digitonin concentration to use for density shift by dose-response assay as follows: Prepare two sets of samples (100 µl/sample) each containing identical amounts of resuspended microsomes. To each set add 270 µl of appropriate digitonin solution to give final concentrations of 0 to 9 mg/ml. To one set, also add Triton X-100 to 0.2% final concentration; to the other set, add equal amounts of water. Incubate 15 min at 0°C.

   Be sure to add the digitonin in the same final volume and as a solution that contains 0.25 M sucrose. Dilute the 20 mg/ml digitonin stock solution to 15 mg/ml using 1 M sucrose; then mix 0.25 M sucrose with 15 mg/ml digitonin in sucrose to give 1.5, 3, 4.5, 6, and 12 mg/ml digitonin, corresponding to final concentrations of 1.1, 2.2, 3.3, 4.4, and 8.8 mg/ml digitonin respectively after addition to samples.

   The Triton X-100 treatment solubilizes the membrane; this set of samples is used to assay total activity. The set of samples treated with water is used to determine latency of activity—i.e., only limited activity will be detected until digitonin reaches a concentration that causes membrane lysis, at which point nearly the same activity will be detected as in the presence of Triton.

5. Assay the samples for the marker activities of specific organelle(s)—i.e., for the presence of an enzyme activity or other marker that is located in the interior of the vesicle of interest. Determine the highest digitonin concentration that does not cause membrane lysis.

   In the dose-response performed by Amar-Costasec et al. (1974), 2.7 vol digitonin-containing solution was added to each sample, and the samples were assayed for ER nucleoside diphosphatase and Golgi galactosyltransferase activities. The appropriate sublytic concentration was determined to be 1.7 mg/ml digitonin (Merck), which the authors estimate corresponds to a molar ratio of digitonin/cholesterol of slightly greater than 1:1.

Prepare sucrose density gradients

6. In the bottom of each centrifuge tube, place a cushion of 2.7 M sucrose.

   For an SW28 tube with 37-ml capacity, this cushion is 3 ml.

7. Using the gradient-forming device and working at room temperature, generate a linear sucrose gradient above the cushion extending from 1.9 M sucrose (density 1.25 g/ml) to 0.8 M sucrose (density 1.10 g/ml) as follows: place the 0.8 M sucrose gradient stock in the back chamber of the gradient maker (15 ml for SW28 tube) and allow it to flow into the junction between the two chambers. Place 15 ml of 1.9 M sucrose gradient stock in the mixing chamber. Open the junction between the chambers of the gradient maker, turn on the stirring device in the mixing chamber, and activate the peristaltic pump to generate the gradient. Once the gradient is generated, chill in a 4°C cold room.

   Because of the high molarity and viscosity of the 1.9 M sucrose, it is placed in the mixing chamber and the gradient is generated high density first. Thus the outlet tube must remain above the surface of the gradient as it is formed in the centrifuge tube (see Fig. 4.1.2). The gradients are prepared at room temperature to decrease viscosity.

Treat microsomes and subject to density gradient centrifugation

8. Prepare two samples from the same batch of microsomes used for the dose-response assay (steps 4 and 5)—1 ml each, if an SW28 rotor will be used, or smaller if scaled down—in homogenization medium. Add 2.7 vol of appropriate-concentration ice-cold digitonin solution to one of the samples. Incubate 15 min on ice.

CU IASI (CENTRAL UNIVERSITY LIBRARY)

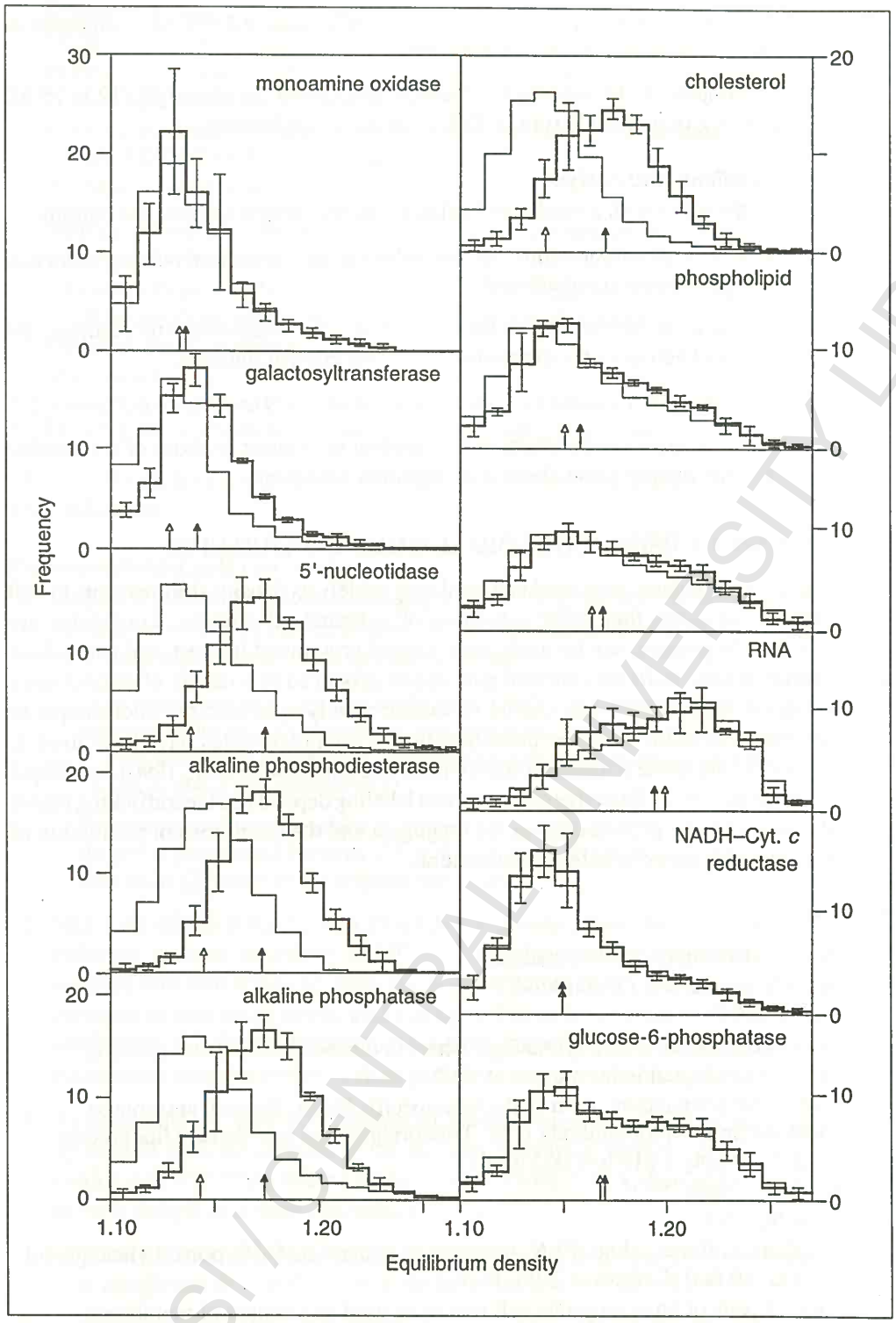


Figure 4.2.7 Distribution of marker activities for various organelles after equilibrium centrifugation. Thin lines and open vertical arrows show the distribution and median density, respectively, of control samples; bold lines and filled vertical arrows show the distribution and median density, respectively, of digitonin-treated samples. Activity profiles are determined across the entire gradient (1.10 to 1.27 g/ml); digitonin treatment shifts the activity profile to higher density in proportion to the amount of cholesterol in the membrane. ER (glucose-6-phosphatase; NADH–cytochrome *c* reductase) and mitochondrial (monoamine oxidase; NADH–cytochrome *c* reductase) marker activities are practically unshifted, corresponding to the negligible cholesterol content of these organelles. Golgi (galactosyltransferase) and plasma membrane (5′-nucleotidase, alkaline phosphodiesterase, and alkaline phosphatase) marker activities are shifted by 0.015 and 0.030 to 0.036 g/ml (median density), respectively. From Amar-Costasec et al. (1974) by permission of The Rockefeller University Press.

DEAKIN UNIVERSITY / CENTRAL LIBRARY

For liver microsomes, Amar-Costasec et al. (1974) used 0.235% (w/v) digitonin in 0.25 M sucrose to give 1.7 mg/ml digitonin.

9. Load the samples on the continuous gradients and centrifuge overnight (12 to 15 hr) at ~82,500 × *g* (e.g., 26,000 rpm in SW28) to reach equilibrium.

Collect the gradients and analyze

10. Puncture the tubes with a needle and collect fractions dropwise from the bottom.

Because the high concentration sucrose solutions are so viscous, manual collection with a micropipettor is not advised.

11. Assay the fractions for the marker activities of specific organelles and compare the distributions of activities for digitonin-treated and control samples.

Examples of results obtained by Amar-Costasec et al. (1974) are shown in Figure 4.2.7.

12. If relevant, compare the distribution of a protein of interest to those of the marker proteins in the presence and absence of digitonin treatment.

BASIC PROTOCOL 8

DENSITY SHIFT USING COLLOIDAL GOLD CONJUGATES

Colloidal gold conjugates have not been used very widely as density shift reagents in cell fractionation; however, they offer a number of potential advantages. Conjugates are relatively easy to prepare, can be made with various proteins of interest, and can induce a substantial density shift; the colloidal gold can be produced in a variety of distinct sizes and, because it scatters electrons, can be visualized readily in an electron microscope. In the following procedure, different protein–colloidal gold conjugates have been used to prepare specific organelle fractions from lymphocytes (Beaumelle et al., 1990) and HEp.2 cells (Futter et al., 1995). Selective compartment labeling depends on the trafficking routes normally taken by the protein used in the conjugate and the conditions of incubation of the conjugate with the cells before fractionation.

Materials

1% (w/v) trisodium citrate dihydrate
1% (w/v) tannic acid (Mallinkrodt)
25 mM K_2CO_3
1% (w/v) chloroauric acid ($HAuCl_4$; Fisher) in twice-distilled water (store in brown bottle; stable for months at 4°C)
Proteins for conjugation: ricin (galactose-specific lectin; Sigma), anti-mouse transferrin receptor antibody (I.W. Trowbridge), and low-density lipoprotein (LDL; density 1.019 to 1.063 g/ml)
10% (w/v) NaCl
1% (w/v) BSA
Phosphate-buffered saline (PBS, *APPENDIX 2E*) containing 0.5% polyethylene glycol (PEG) 20,000 (Carbowax 20M, Fluka)
Cultured cells of an appropriate cell line to be used as a source of membrane fractions: e.g., BW 5147 mouse T cell lymphoma cells (characterized by I.W. Trowbridge, Salk Institute; available from ATCC) or HEp.2 cells (ATCC CCL23)
Dulbeccos Modified Eagle Medium (DMEM) containing 100 mg/ml BSA
DMEM containing 1% horse serum
DMEM containing 5 mg/ml BSA
PBS (*APPENDIX 2E*)
Hypoosmotic homogenization medium (see recipe)
Hyperosmotic buffer (see recipe)

CU IASI CENTRAL UNIVERSITY LIBRARY

RNase (Sigma)
DNase I for treating cell homogenates (Sigma)
Pellet resuspension solution: 0.25 M sucrose (ultrapure; e.g., ICN Biochemicals)/1 mM EDTA/10 mM Tris acetate (pH 8.0)
Continuous density gradient solutions: 30% and 60% (w/v) sucrose (ultrapure) in 1 mM EDTA/10 mM Tris acetate (pH 7.5)
2.5 M sucrose in water

Nitrogen bomb and Dounce homogenizer for homogenization
Low-speed refrigerated centrifuge
Ultracentrifuge with fixed-angle and swinging-bucket rotors (e.g., Beckman ultracentrifuge with Type 50 and SW40 or SW41 rotors)
Gradient-forming device
Micropipettor

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after harvesting cells should be conducted at 0°C (ice bucket) or 4°C (cold room).

Prepare colloidal gold conjugates

1. Prepare a reducing mixture containing 4 ml of 1% trisodium citrate dihydrate, 0 to 5 ml of 1% tannic acid (depending on the particle size desired), and the same volume of 25 mM K_2CO_3. Add Milli-Q water to 20 ml.

   As discussed in detail by Slot and Geuze (1985), the amount of tannic acid added determines the size of the particles obtained in the gold sol. Addition of 0.01, 0.1, 0.5, or 5 ml tannic acid yields (after the following steps) gold particles of average diameter 16, 10, 5, and 3 nm, respectively.

   Potassium carbonate is added to compensate for the acid from the tannic acid so that the pH is maintained between 7.5 and 8. Slot and Geuze used glass-distilled (rather than Milli-Q) water in the original procedure.

2. Mix 1 ml of 1% $HAuCl_4$ with 79 ml Milli-Q water. Heat the gold solution and the reducing mixture separately to 60°C, then rapidly mix together with stirring. The solution will turn wine red when the gold sol forms (this takes a few seconds in the presence of maximum tannic acid and up to 1 hr in the absence of tannic acid). After sol formation is finished, continue to heat the samples until boiling. Allow to cool to room temperature and store in clean glass containers until used (up to several months).

3. To determine the minimum amount of protein required to stabilize the gold sol, set up several 0.25-ml samples of sol and add increasing amounts of the protein to be conjugated (0 to 10 µg using 2 mg/ml stocks in water). Mix, then add 10% (w/v) NaCl to each sample to a final concentration of 1%.

   If the solution turns blue, then the gold sol has flocculated and the amount of protein is insufficient for stabilization; if the solution stays red, the sol is stable.

   Slot and Geuze (1985) state that blue color development is slow and subtle in samples that retain excess tannic acid. In this case they recommend adding 0.1% to 0.2% H_2O_2 to destroy the tannic acid and facilitate normal color development.

   Ricin, anti-mouse transferrin receptor antibody, and LDL are the proteins conjugated in the work of Beaumelle et al. (1990), which is being used as an example. Ricin is a galactose-specific lectin (CAUTION: *Ricin is a toxin and must be handled with care). Anti-mouse transferrin receptor antibody is discussed in Lesley et al. (1984). Low-density lipoprotein (LDL; density 1.019 to 1.063 g/ml) is purified from fresh human plasma*

BOSTON UNIVERSITY MEDICAL LIBRARY

using the procedure of Havel et al. (1955), which involves flotation from NaCl/KBr solutions.

4. Prepare a supply of the stable conjugate by adding more than twice the minimum amount of stabilizing protein with stirring. After 1 min, add BSA to 0.1% final (for added stability).

   Protein-gold conjugates formed this way are quite uniformly sized.

   Beaumelle et al. (1990) conjugated 1.9 mg ricin and 1.0 mg anti-transferrin receptor antibody with 100 ml gold.

5. Pellet the protein-gold conjugates in a fixed-angle rotor for 45 min at 50,000 × *g* (for gold ≥10 nm) *or* at 125,000 × *g* (for gold <10 nm), 4°C. The pellets usually have a tightly packed portion that adheres to the tube and a larger loose portion that is easily withdrawn with a pipet. Take only the latter.

6. Resuspend the protein-gold conjugate in 3 to 5 ml PBS containing 0.05% PEG 20,000 (as an extra stabilizer).

   Where protein-gold conjugates are made with serum LDL, the conjugation procedure used is different (Handley et al., 1981), and the conjugate is used immediately. A solution of 1.5 µg LDL in 0.5 ml of 50 mM EDTA (pH 5.5) is added with rapid stirring to 5.0 ml colloidal gold (prepared at the same pH using sodium citrate). The LDL-gold conjugate is sedimented away from excess LDL through a 35% sucrose cushion (centrifuging 20 min at 9000 × g) in a swinging-bucket rotor. The loose pellet is resuspended in cell culture medium and used directly for incubation.

Incubate protein-gold conjugates with cells

7a. *For isolation of plasma membranes*: Incubate 10^8 cultured cells 1 hr on ice in 5 ml DMEM containing 100 µg/ml BSA and 200 µl gold-ricin conjugate (from step 6).

   This step stops all membrane traffic. Beaumelle et al. (1990) report that the lectin concanavalin A does not bind to the plasma membrane sufficiently to cause a density shift during subsequent fractionation.

7b. *For isolation of endosomes*: Incubate 10^8 cultured cells 2 hr at 22°C in 5 ml DMEM containing 100 µg/ml BSA and 130 µl anti-mouse transferrin receptor antibody–gold conjugate.

   Incubation at 22°C retards both recycling back to the cell surface and trafficking from endosomes to lysosomes. Note that under steady-state conditions, 80% of transferrin receptors in these cells are concentrated in endosomes.

7c. *For isolation of coated pits*: Incubate 10^8 cultured cells 2 hr at 10°C in 5 ml DMEM containing 100 µg/ml BSA and 130 µl anti-mouse transferring receptor antibody–gold conjugate.

   Incubation at 10°C inhibits internalization of coated pits. Note that a large fraction of the transferrin receptor that is present on the cell surface is concentrated in coated pits.

7d. *For isolation of lysosomes*: Incubate 2 to 5 × 10^8 cultured cells 18 hr at 37°C in DMEM containing 1% horse serum, 5 mg/ml BSA, and 0.8 mg LDL-gold conjugate. Withdraw the medium with a pipet and replace with DMEM containing 5 mg/ml BSA and 150 µg/ml LDL. Incubate an additional 3 hr at 37°C.

   The second incubation (with unconjugated LDL) serves to chase internalized LDL to lysosomes and compete away any surface-associated LDL-gold.

8. After incubation, wash the cells twice with DMEM containing 100 µg/ml BSA and once with PBS (both at the temperature used in step 7a-d), then chill on ice.

CCU IASI / CENTRAL UNIVERSITY LIBRARY

Lyse and homogenize labeled cells and prepare postnuclear supernatant

9. Wash the cells once with hypoosmotic homogenization medium, then scrape (or resuspend, if not attached) into the same buffer, using 1 ml buffer per 5×10^7 cells.

10. Place the cells in a nitrogen bomb for 20 min at 750 lb/in^2 and release pressure according to manufacturer's instructions. Transfer the initial homogenate to a Dounce homogenizer (without altering the composition of the medium) and homogenize with 20 strokes.

    The initial homogenization using the nitrogen bomb results in ~30% lysis of the BW 5147 cells (judged by trypan blue exclusion; UNIT 5.5). After Dounce homogenization lysis is ~90% with no increase in nuclear damage.

    If a nitrogen bomb is not available, then other homogenization procedures should be tried. Other publications from the same laboratory that developed this procedure report homogenization by passage 15 times through a syringe equipped with a 21-G needle (Futter and Hopkins, 1989) and by repeated passage through a micropipet tip (Beardmore et al., 1987). To a large extent, the homogenization procedure that works best will depend on the type of cells being studied.

11. Add 0.1 vol hyperosmotic buffer and centrifuge 5 min at 800 × *g*, 4°C.

12. Remove the supernatant (PNS). Add 0.1 mg/ml RNase and 0.2 mg/ml DNase I and incubate 3 min at 37°C.

    This step digests any accessible chromatin and RNA and improves the ensuing purification by decreasing aggregation of the total membrane pellet that is prepared in the next step.

Pellet the total membranes of the PNS

13. Pellet organelles in the PNS by centrifuging 30 min at 120,000 × *g*, 4°C, in a fixed-angle rotor.

    These centrifugation conditions will reliably pellet organelles that contain bound colloidal gold.

14. Resuspend the pellets in pellet resuspension solution, using a small-volume Dounce homogenizer if necessary to obtain thorough dispersal.

Prepare and run the continuous sucrose gradients

15. Place sucrose cushions consisting of 0.5 ml of 2.5 M sucrose in the bottom of 13-ml ultracentrifuge tubes.

    Instructions are given for 13-ml tubes and a SW40 or SW41 rotor, but may be adjusted to different sizes if desired.

16. In each tube, generate a 12-ml continuous (60% to 30%) sucrose gradient above the cushion using the gradient solutions as follows: place 6 ml of 30% sucrose solution in the back chamber of the gradient maker, and fill the junction tube to the mixing chamber. Place 6 ml of 60% sucrose solution in the mixing chamber. Open the junction between chambers, start stirring and pumping, and generate gradients (high density first) above the cushion.

    The delivery tube from the gradient-forming device should be kept above the surface of the fluid in the centrifuge tube.

17. Load 0.5-ml aliquots of the resuspended pellet on top of each gradient and centrifuge 16 hr at 200,000 × *g*, 4°C, in an ultracentrifuge with swinging-bucket rotor (e.g., SW40 or SW41).

BOWLING GREEN STATE UNIVERSITY LIBRARY

Collect the fractions for analysis

18. Unload the gradient from the top in 1-ml fractions. Manual collection with a micropipettor will work fine as long as the high-viscosity sucrose is collected slowly. Otherwise use an automated collection device (Buchler, Nyegaard, or MSE Scientific; see Alternate Protocol 1, step 11).

    In all cases the organelles that have been purified by density shift are found in the fraction that is just above the pinkish pellet.

19. Analyze the fractions by assaying for marker activities or for protein by two-dimensional SDS-PAGE (*UNIT 10.4*), electron microscopy, and immunoprecipitation or immunoblotting.

    In the published procedures, the authors also carried out all these procedures using ^{125}I-radiolabeled protein-gold conjugates and were able to estimate yields of the various organelles by the density shift technique. They obtained 40%, 32%, 10%, and 16% of the associated radiolabel in the high-density fractions of plasma membranes, endosomes, coated pits, and lysosomes, respectively.

BASIC PROTOCOL 9

EXTRACTION OF EXTRINSIC PROTEINS FROM MEMBRANES USING SODIUM CARBONATE

The methods for cell fractionation presented elsewhere in this unit focus mostly on purifying specific organelles without further subfractionation into component parts. For many studies that focus on membranes of organelles or on specific membrane-associated proteins, it is desirable to have an easy way to categorize integral and extrinsic polypeptides. Extraction with 0.1 M sodium carbonate (pH 11.5) at 0°C (Fujiki et al., 1982) has gained widespread popularity as an easy and efficient method for selectively stripping extrinsic proteins off membranes without affecting the disposition of integral components (including transmembrane and lipid-anchored proteins). The treatment converts vesicles into small membrane sheets, releasing soluble and adsorbed proteins in the process. The membranes are recovered after the treatment by centrifugation. Carbonate extraction has been applied to numerous organelles including endoplasmic reticulum, Golgi, secretory vesicles, plasma membranes, lysosomes, and peroxisomes.

Materials

0.1 M sodium carbonate (Na_2CO_3), pH 11.5, ice cold
0.3 M sucrose (ultrapure; e.g., ICN Biochemicals), buffered to pH 7 to 7.5 (e.g., with 10 mM Tris·Cl)
Cell organelle or membrane fraction to be treated
Ultracentrifuge with swinging-bucket rotor

NOTE: Prepare solutions with Milli-Q-purified water or equivalent.

NOTE: All operations after removing and weighing tissue should be conducted at 0°C (ice bucket) or 4°C (cold room).

1. Pellet the organelles or vesicles to be extracted by centrifugation, using conditions described in the appropriate basic or alternate protocol.
2. Resuspend the pellets in a small amount of buffered sucrose sufficient to obtain a well-dispersed suspension.
3. Dilute the sample 50× to 1000× with ice-cold 0.1 M Na_2CO_3. Maintain 30 min on ice.

    A final protein concentration of 10 to 1000 μg/ml is generally appropriate. Assay protein content if there is concern that this range is exceeded.

CU IAST / CENTRAL UNIVERSITY LIBRARY

4. Place a small cushion of 0.3 M buffered sucrose in a centrifuge tube (~10% of the volume of the tube), then carefully layer the organelle suspension on top. Pellet the membranes by centrifuging 1 to 2 hr at ~150,000 × *g*, 4°C.

   Although not specified in the original procedure, the buffered sucrose cushion prevents formation of a rock-like pellet that is often very difficult to resuspend.

5. Resuspend the pellets and use for assays.

   Although the high-pH treatment on ice generally does not affect subsequent immunoprecipitations, it can inhibit some enzyme activities (check in a preliminary experiment).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Digitonin density shift gradient stock solutions

Prepare three separate stock solutions:

0.8 M sucrose/3 mM imidazole, pH 7.4 (density 1.10 g/ml)
1.9 M sucrose/3 mM imidazole, pH 7.4 (density 1.25 g/ml)
2.7 M sucrose (density 1.34 g/ml)

The 2.7 M sucrose is used for the cushion and contains no buffer. Sucrose should be ultrapure (e.g., ICN Biochemicals).

Sucrose stocks can be filtered using 1.2-μm Millipore filter (or equivalent) and stored at least 2 weeks at 4°C or indefinitely at –20°C. 1 M imidazole (pH adjusted with concentrated HCl) can be filtered using 0.22-μm Millipore filter (or equivalent) and stored indefinitely at 4°C.

Digitonin stock solution, 20 mg/ml

Dissolve 20 mg/ml digitonin in water, stirring for 15 min at 90° to 95°C. Filter through 0.22-μm Millipore filter (or equivalent) and store up to several weeks at 4°C.

Digitonin is available from Merck and Sigma. Neither preparation is pure (they are contaminated mostly by other saponins that bind to cholesterol), but for purposes of the density shift experiment they can be treated as if pure. However, it is preferable not to switch brands without retitration.

Homogenization medium (lectin adsorption)

0.28 M sucrose (ultrapure; e.g., ICN Biochemicals)
50 mM 2-*N*-morpholinoethanesulfonic acid (MES), pH 6.0
150 mM NaCl
10 mM $MgCl_2$
0.1 mg/ml soybean trypsin inhibitor (Sigma)
0.1 mM phenylmethanesulfonyl fluoride (PMSF; Sigma)

PMSF has a very short half-life in aqueous solutions and must be added from a 0.2 M stock in ethanol (stored at –20°C) just before fractionation. Medium without inhibitors can be stored indefinitely at –20°C.

Pancreatic zymogen granules are labile above pH 7 and the low pH is used to keep them as intact as possible; 150 mM NaCl is included because it is reported to promote plasma membrane vesiculation with ectodomain facing outwards.

Homogenization medium (velocity differential centrifugation)

0.25 to 0.3 M sucrose (ultrapure; e.g., ICN Biochemicals)
50 mM Tris·Cl (pH 7.5; optional)
25 mM KCl (optional)

continued

BOLDRINI CENTRAL UNIVERSITY LIBRARY

5 mM $MgCl_2$ (optional)
Proteinase inhibitor cocktail (optional)

Although sucrose is the critical component of this medium, other additives are often included as well. Generally the medium is prepared fresh before use.

A suitable proteinase inhibitor cocktail includes 1 µg/ml each (final concentration in sucrose homogenization medium) of leupeptin, antipain, pepstatin, and aprotinin (some investigators also include o-phenanthroline; all available from Sigma), along with 0.2 mM AEBSF (Calbiochem) or PMSF (phenylmethanesulfonyl fluoride), in 1 mM EDTA (pH 7). These components are added from 1000× stock solutions prepared ethanol (pepstatin and PMSF) or water and stored at –20°C. PMSF has a very short half-life in aqueous solution and should be added just before starting the homogenization.

Hyperosmotic buffer

700 mM KCl
40 mM magnesium acetate
1 mM dithiothreitol (DTT)
10 mM HEPES, pH 7.5
Make fresh before use

Store DTT at –20°C as a 1 M stock solution in water.

Hypoosmotic homogenization medium (colloidal gold density shift)

15 mM KCl
1.5 mM magnesium acetate
1 mM DTT
10 mM HEPES, pH 7.5
Make fresh before use

Store DTT at –20°C as a 1 M stock solution in water.

Percoll gradient stock solutions

No-Percoll solution (low sucrose)
0.3 M sucrose
5 mM MOPS or MES buffer
1 mM EDTA
0.2 µg/ml DPPD (using 0.4 mg/ml DPPD in ethanol)
Adjust pH to 6.5-6.7 using 1 N HCl

No-Percoll solution (high sucrose)
2.15 M sucrose
35 mM MOPS or MES
7 mM EDTA

Prepare solution from the following stock solutions: 2.5 M sucrose (ultrapure; e.g., ICN Biochemicals); 1 M 2-N-morpholinoethanesulfonic acid (MES), pH 6.5; 1 M 3-N-morpholinopropanesulfonic acid (MOPS), pH 6.8; 0.2 M EDTA, pH 7.0; and 0.4 mM DPPD (diphenyl-p-phenylenediamine; Kodak) in ethanol. 2.5 M sucrose can be stored indefinitely at –20°C; 1 M MES or MOPS and 0.2 M EDTA should be filtered using 0.22-µm Millipore filter (or equivalent) and can then be stored indefinitely at 4°C; and DPPD stock should be made fresh before use.

The pK_a values of MOPS and MES are 7.2 and 6.1, respectively.

60% Percoll/sucrose: Mix 60 ml Percoll, 14 ml No-Percoll solution (high sucrose), and 26 ml water. Ajust pH as described above and add 0.2 µg/ml DPPD (from 0.4 mg/ml DPPD stock in ethanol).

86% Percoll/sucrose: Mix 86 ml Percoll and 14 ml No-Percoll solution (high sucrose). Adjust pH as described above and add 0.2 µg/ml DPPD (from 0.4 mg/ml DPPD stock in ethanol).

Percoll is slightly alkaline, so it is important to check and adjust the pH, especially if the organelles to be purified are labile at alkaline pH.

PCU IASRI CENTRAL UNIVERSITY LIBRARY

Rate zonal centrifugation gradient stock solutions

Low-density gradient solution	*High-density gradient solution*
0.5 M sucrose	1.8 M sucrose
4% (w/v) Ficoll 400	4% (w/v) Ficoll 400
5 mM MES ($pK_a = 6.1$)	5 mM MES
1 mM EDTA	1 mM EDTA
0.2 µg/ml DPPD	0.2 µg/ml DPPD

Prepare gradient solutions from the following stock solutions: 2.5 M sucrose (ultrapure; e.g., ICN Biochemicals); 50% Ficoll 400; 1 M 2-N-morpholinoethanesulfonic acid (MES), pH 6.5; 0.2 M EDTA, pH 7.0; and 0.4 mg/ml DPPD (diphenyl-p-phenylenediamine; Kodak) in ethanol. Sucrose and Ficoll stocks are stored long-term at –20°C; gradient solutions are used within two weeks of preparation; and the DPPD stock is made fresh and used without storage.

Running buffer

0.8 M sorbitol
10 mM triethanolamine
1 mM EDTA, pH adjusted to 7.2 with acetic acid (TEA)

Sorbitol is used because yeast metabolize sucrose. The running buffer generally is prepared fresh before use.

COMMENTARY

Background Information

Organelle purification procedures capitalize on the differences in size, density, and (occasionally) surface charge density of individual types of organelles. Most fractionation procedures that are based on centrifugation involve some combination of procedures that distinguish both size and density. Initially, a homogenate is prepared in isoosmotic (or slightly hyperosmotic) sucrose or some other predominantly nonelectrolyte medium. Sucrose solution deviates from ideality; while 0.3 M sucrose and 150 mM NaCl have equivalent particle concentrations, 0.25 M sucrose and 150 mM NaCl have the same osmotic activities. Investigators generally homogenize cells and tissues in solutions that are somewhere in the range of 0.25 to 0.3 M sucrose. The homogenate is filtered through cheese cloth or nylon screen to remove residual large connective tissue debris. Then differential centrifugation by velocity is used at low speed to pellet large particulates—nuclei, erythrocytes, and larger particles (including unbroken and partially disrupted cells). Subsequent treatment of the postnuclear supernatant (PNS) is a function of the mission of the experiment as discussed in the protocols.

Differential centrifugation by velocity (Basic Protocol 1) separates organelles within a cell homogenate, largely on the basis of organelle size, through successive spins of increasing centrifugal field and time in a 0.25 to 0.3 M sucrose medium. What does not sediment under the first set of conditions may sediment under one of the later sets of conditions. As all cellular organelles have a greater density than the medium, the separation is largely based on organelle size. The pellets produced during each centrifugation are enriched for organelles of a certain size but are also substantially contaminated with smaller organelles. DeDuve and Berthet (1954) calculated that in a tube where the bottom is twice as far as the surface of the liquid from the axis of rotation, a pellet containing the fastest sedimenting organelle will be contaminated by 40% of an organelle that sediments at one-third the speed and by 13% of an organelle that sediments with one-tenth the speed. Generally, further enrichment is necessary, and purity can be improved by resuspending the selected pellet in medium to its original volume and recentrifuging under the same conditions.

Although differential centrifugation by velocity is mainly used to separate and concentrate crude organelle fractions, differential centrifugation by equilibrium (Basic Protocol 2) is generally used as a follow-up to further purify organelles from contaminants according to buoyant density differences. Continuous density gradients are widely used to generate unique (but often partially overlapping) distributions along the gradient for various types of organelles within a complex mixture. The resolution is well suited for analytical purposes

POLYTECHNIC UNIVERSITY LIBRARY

where the goal is not necessarily preparative purification. In contrast, discontinuous (step) gradients are most frequently used for preparative purifications. Densities of media are selected so that the organelle of interest concentrates at the interface between particular density steps while other organelles either sediment or float to other interfaces.

An alternative to differential separation by velocity or equilibrium is rate zonal centrifugation (Alternate Protocol 1), a nonequilibrium method in which organelles separate based on a combination of size, density, and shape into enriched regions within the gradient. This method is particularly suitable for separating several types of organelles into distinct distributions within a relatively short time (1 to 3 hr rather than 8 hr to overnight as in equilibrium centrifugation). However, it should not be viewed as a substitute for differential centrifugation by equilibrium, which is one of the best ways to separate fractions of organelles of similar size but different density (e.g., rough and smooth microsomes).

Percoll-gradient fractionation (Alternate Protocol 2) has several advantages. The density gradients are isoosmotic with the homogenization medium, and due to the low viscosity, organelles approach equilibrium density quite rapidly (30 min for 1-μm organelles like granules and mitochondria; 2 hr for small (≤0.2-0.3 μm) vesicles like microsomes and endosomes). Also shallow density gradients are generated with relatively brief centrifugation; this can be advantageous for obtaining good spatial separation of organelle populations that differ only slightly in density from one another. Note, however, that because the shallow region of the gradient is bounded by regions where density changes sharply with distance in the tube (see Fig. 4.2.2), organelles with similar densities but not separated across the shallow region will tend to pile up at the boundaries. Thus Percoll is especially advantageous when the experimental goal can be reduced to separating two components from one another (or one component from the rest), and it is unlikely to be the method of choice for resolving complex mixtures into multiple subfractions that are individually enriched in different organelles. Also, it should be kept in mind that if a particular type of organelle has a heterogeneous density (e.g., endosomes), portions of the total population may be distributed on both sides of the shallow resolving segment of the gradient, depending on the concentration of Percoll selected for use.

Deuterium oxide fractionation (Alternate Protocol 3) is an equilibrium density procedure in which H_2O is partly replaced with D_2O (heavy water). It is based on the fact that the density of D_2O is 1.1 g/ml—substantially closer than that of water to the density of biological organelles. In principle, this approach should be quite useful as an aid in separating organelles, for two reasons. First, inclusion of D_2O in density gradients can significantly reduce the osmolarity of solutions that are isopycnic with most biological organelles. Second, replacement of H_2O with D_2O can reduce the density *difference* between organelles and the medium. For two populations of organelles that differ only slightly from one another in density and sediment at only slightly different rates in H_2O-based media, reducing the density difference between the organelles and the medium is one way to magnify the differences in sedimentation rate (this can be verified using the Svedberg equation—see *UNIT 4.1*). In practice, the practical value of D_2O has fallen short of expectations, mostly because the densities of several organelles increase in heavy water due to exchange of D_2O for bound water. However, such exchange may be advantageous if it occurs to differing extents between organelles and thereby magnifies their density differences. For example, Beaufay et al. (1959) showed that the buoyant density of liver mitochondria is largely unchanged in H_2O/sucrose vs. D_2O/sucrose, whereas the buoyant densities of both lysosomes and peroxisomes substantially increased in the presence of heavy water.

Both D_2O and glycerol (another popular density-modifying agent, e.g., Wagner et al., 1978) readily permeate membranes. In these media, internal and external solvents are identical. For organelles with relatively little internal solute (e.g., synaptic vesicles), sedimentation in D_2O or glycerol gradients has been quite useful in analyzing the physical characteristics of the membranes with minimized contributions from the internal solvent (Wagner et al., 1978).

Excellent fractionation procedures for a variety of cellular organelles have been described that make use of density gradient centrifugation in nonionic iodinated solutes (Alternate Protocol 4). As examples: Barlowe et al. (1994) report use of isopycnic centrifugation in Nycodenz for isolation of biosynthetically labeled, ER-derived transport vesicles from a cell-free assay; Loh et al. (1984) use metrizamide gradients in sucrose for purifying storage granules for peptide hormones from intermediate pituitary, demonstrating reasonably good

separation of granules from lysosomes; and Wattiaux et al. (1978) use metrizamide gradients to prepare highly purified lysosomes, well resolved from mitochondria and peroxisomes. The last two studies thoroughly account for a variety of marker enzyme activities.

Gel filtration (Basic Protocol 3) using controlled-pore glass (CPG-3000, Sigma; mean pore diameter 3000 Å) has been employed routinely to prepare highly purified synaptic vesicles from rat brain (Huttner et al., 1983). Sephacryl S-1000 (Pharmacia Biotech), which also excludes particles larger than 3000 to 4000 Å in diameter, has been used to purify secretory vesicles, clathrin-coated vesicles and ER-Golgi/nonclathrin-coated transport vesicles from the yeast *S. cerevisiae* (Walworth and Novick, 1987; Mueller and Branton, 1984; and Barlowe et al., 1994).

For complex mixtures of small vesicles, differences in size and density between individual types of vesicles are frequently insufficient to enable their separation as purified populations. In some of these cases, electrophoretic purification techniques—which capitalize on differences in surface charge density among the organelles—have proved useful at both the preparative and the analytical levels.

Organelle purification using preparative electrophoresis (exemplified by Basic Protocol 4) has been performed in large-porosity solid matrices such as agarose—e.g., for clathrin-coated vesicles from bovine brain and liver and from CHO cells (Rubenstein et al., 1981) and the ribonucleoprotein particles known as vaults from rat liver (Kedersha and Rome, 1986). It has also has been carried out in liquid using density gradients (which provide stability but do not contribute to the separation). The latter procedure has been used in the preparative purification of endosomes from cultured cells (Marsh et al., 1987) and secretory vesicles from *S. cerevisiae* mutants (Holcomb et al., 1987, and Basic Protocol 4). In each case, preenrichment of the samples using centrifugation techniques have been used to overcome the limitations in total protein that can be loaded.

More commonly, electrophoresis has been used at the analytical level to separate and analyze organelles that are selectively labeled by markers, as exemplified by Alternate Protocol 5. Cases where analytical electrophoresis has been applied include synaptic vesicles isolated from electric organs of marine elasmobranchs (Carlson et al., 1978) and specialized endosomes that accumulate MHC class II molecules (Amigorena et al., 1994; Tulp et al., 1994, and Alternate Protocol 5).

Both preparative electrophoresis and analytical density-gradient electrophoresis capitalize on equipment that is relatively inexpensive. In the event that a free-flow electrophoresis apparatus (e.g., Bender and Hobein Elphorvap II; cost $100,000) is available, then the procedure for its use presented in detail in Marsh et al. (1987) and subsequently employed by Amigorena et al. (1994) should be seriously considered for either preparative- or analytical-scale isolations.

Isolation of cellular organelles by solid-phase immunoadsorption using specific antibodies (Basic Protocol 5) has become very popular in the past several years. Clearly, use of these procedures requires availability of a specific antibody capable of reacting with an antigenic epitope exposed on the surface of the organelle of interest, and relies on the quality of that primary antibody and the extent to which the antigen is concentrated in the compartment of interest. These procedures can be conducted at both the preparative and analytical levels. Earlier procedures used either fixed *S. aureus* or protein A–Sepharose with adsorbed antibody (e.g., Merisko et al., 1981). This approach has continued to be used, especially for preparative-level procedures (e.g., isolation of glucose transporter vesicles; Laurie et al., 1993). As an alternative (as described in Basic Protocol 5), magnetic beads that are coated with secondary antibodies (e.g., sheep anti-mouse Fc or sheep anti-rabbit Fc) provide a new-generation solid support that allows easy isolation and is especially amenable to both functional and morphological analysis of associated organelles. The use of magnetic beads as a solid support for isolating endosomes from cultured cells and subsequently analyzing their fusion has been elegantly demonstrated by Gruenberg and Howell (1986). The properties of commercially available magnetic beads are discussed in Howell et al. (1989).

Basic Protocol 6 presents a method that uses affinity chromatography on the lectin wheat germ agglutinin (WGA) to purify apical plasma membranes from homogenates of pancreatic acinar cells (Paul et al., 1992; *UNIT 9.2*). WGA that binds both *N*-acetylglucosamine (GlcNAC) and *N*-acetyl neuraminic acid residues in glycoproteins, and interactions with WGA can be competed with free GlcNAc, which is relatively inexpensive. Other examples of lectin adsorption in subcellular frac-

BARNARD COLLEGE LIBRARY

tionation include use of WGA conjugated to iron dextran for purification of plasma membrane vesicles from CHO cells in a strong electromagnetic field (Warnock et al., 1993) and WGA–colloidal gold conjugates to remove plasma membranes prior to purification of other organelles (Gupta and Tartakoff, 1989).

A variety of density-shift techniques exist that are employed for both analytical and preparative purposes. Basic Protocol 6 presents a particularly useful density-shift strategy in which digitonin is applied to a crude organellar fraction; the compound binds to cholesterol, thereby increasing the buoyant density of different organelles to a lesser or greater extent depending on the cholesterol content of the organellar membranes. The protocol is based on the approach of Amar-Costasec et al. (1974), who applied it to a crude microsomal fraction from rat liver to demonstrate a substantial shift of marker activities of the plasma membrane (which has a high cholesterol content), a moderate density shift of a marker activity of the Golgi (which has a lower cholesterol content), and negligible shifts of markers of endoplasmic reticulum (rough microsomes) and mitochondria (which both have little or no membrane cholesterol). In turn, the density shifts of these standard markers can serve as references for comparing the behavior of proteins whose localizations are not known. A complementary procedure involving use of EDTA or pyrophosphate to strip ribosomes from the endoplasmic reticulum can be used to achieve a selective density shift of rough microsomes (Amar-Costasec et al., 1974).

Both binding to the surfaces of intact cells and endocytosis have also been used to achieve density shifts in selected organelles during subsequent fractionation. In one such procedure, receptor-mediated endocytosis of proteins conjugated to horseradish peroxidase (HRPO) has been used to label endocytic compartments which are then shifted to higher buoyant density by carrying out the cytochemical reaction of HRPO using H_2O_2 and diaminobenzidene (Courtoy et al., 1984). Unfortunately, the oxidative cytochemical reaction induces wholesale protein cross-linking within the labeled compartment, significantly limiting the potential applicability of this approach (Ajioka and Kaplan, 1987). Another interesting (though rather specialized) application used density shift by an enzyme cytochemical reaction product to distinguish clathrin-coated vesicles involved in endocytosis from other types of coated vesicles in liver (Helmy et al., 1986). On the other hand, several procedures using protein–colloidal gold (or iron-dextran) conjugates appear to offer both wider application and greater possibilities for further analysis of the isolated fractions (e.g., Warnock et al., 1993; Futter and Hopkins, 1989; Beaumelle et al., 1990; Futter et al., 1995). Among these, the use of ricin-gold, LDL-gold, and transferrin receptor antibody–gold conjugates to isolate selected compartments involved in endocytosis (Beaumelle et al., 1990) is presented as Basic Protocol 8.

Although Basic Protocol 8 is not an organelle isolation protocol, it is the procedure most widely used to extract extrinsic (or adsorbed) proteins from organelle membranes. Not only has it enabled investigators with an interest in integral membrane proteins to efficiently remove soluble contaminating proteins but it has also been used to distinguish peripheral membrane proteins (those desorbed in the presence of sodium carbonate) from integral membrane proteins (those that remain membrane-associated following treatment). As such, it is complementary to phase partitioning in Triton X-114 in which integral, but not peripheral, proteins bind the detergent and sediment in micelles at the detergent cloud point (Bordier, 1981). Note, however, that lipid-anchored (fatty acylated or prenylated) proteins either partially or fully behave as integral membrane proteins in both procedures.

Critical Parameters and Troubleshooting

These procedures can be performed using either tissue from freshly sacrificed animals (stored frozen tissues are not usually satisfactory) or cultured cells. As compared to cultured cells, however, mammalian tissues offer at least three advantages for cell fractionation and purification of various organelles. First, the cells in mammalian tissues generally are highly differentiated and, depending on the functional specialization of the selected cell type, contain large numbers of organelles per cell. Second, the mass of cells available in tissue taken even from a single rat is much larger than is readily available from reasonably large size cultures. Finally, the connective tissue matrix surrounding cells in tissues contributes to shearing during homogenization and is often an aid in efficiently disrupting cells under reasonably gentle homogenization conditions. Of the procedures discussed in *UNIT 4.1*, a glass mortar and motor-driven Teflon pestle rotating at ~1500 to 2000

rpm is used most widely for routine homogenization of tissues. If the investment of connective tissue is extensive (e.g., for parotid salivary gland) making direct homogenization with a Teflon pestle difficult, then a prehomogenization (10 sec at half-maximal power with a Tekmar Tissumizer (or Polytron) inserted directly in the glass mortar) is used before homogenization with the Teflon pestle.

The following requirements are critical for each purification method:

Differential centrifugation by velocity

The concentration of the homogenate should be standardized somewhere in the 10% to 20% (w/v) range; in especially concentrated homogenates, sedimentation properties change and aggregation is exacerbated.

The early steps of differential centrifugation by velocity may be the only chance in a multistep procedure to remove large organelles. Be sure the centrifugation is hard enough to achieve this task even if it means taking a partial loss of the desired smaller organelles. The loss can be reduced by resuspending the pellet, recentrifuging, and then pooling the second supernatant with the original supernatant.

Differential centrifugation by equilibrium

Density must be measured accurately—e.g., a refractometer should be used. In addition, reproducible preparation and collection of gradients is very important.

Upper limits to gradient loads must be identified so that organelles do not aggregate appreciably and instead continue to sediment according to their individual buoyant densities. Deceleration at the end of the centrifugation should be done with the brake in the OFF position so that gradients are minimally disturbed by mixing.

Following homogenization, vesicles deriving from a particular organelle may be heterogeneous in size and density. Further, the densities of different organelles, especially various types of small vesicles (microsomes, Golgi elements including subcompartments, and endosomes) may not differ much from one another. Thus the investigator should not expect separate, discrete bands. Rather, distributions of marker activities across the gradient must be compared.

Rate zonal centrifugation using sucrose

Reproducible preparation, loading, running, and unloading of gradients is essential. Conditions are particularly important where distributions of activities/markers are compared between tubes rather than within the same tube. Use of a refractometer to check the concentrations of solutions used to generate the density gradient is recommended. The refractometer is also useful for determining the density of gradient fractions. If a programmable centrifuge is available, use the same acceleration/deceleration program from run to run; otherwise, establish and adhere to a specific manual protocol. The separation between different organelle-enriched bands can be manipulated empirically to some extent by changing the length of centrifugation or the density limits of the gradient.

Gradient fractionation using Percoll

Because Percoll itself has negligible osmolarity, it must be added to isoomotic media. Almost all membrane-bounded organelles behave as osmometers and will lyse in hypoosmotic media. Although electrolyte-based solutions such as NaCl are generally used in separating populations of cells, organelles tend to aggregate in such media and are generally fractionated in nonelectrolyte-based media.

Close attention should be paid in preparing Percoll-sucrose mixtures so that starting densities of the self-forming gradients are reproducible. This is critical in separating organelles with very similar densities.

Because the density profile of the Percoll gradient continually changes during the run, centrifugation conditions (time and rpm) should be performed reproducibly. Centrifugation times that are too long can actually elevate the cross-contamination between low- and high-density fractions. After fractionation, it is necessary to remove the Percoll from purified fractions, as it can interfere with further processing. The simplest method to accomplish this is via additional low-speed centrifugations (as described in Alternate Protocol 2) or through ultracentrifugation on a step gradient (Zastrow and Castle, 1987).

If, instead, fractions are concentrated by high-speed centrifugation (as in several of the other protocols in this unit), Percoll will also sediment; it forms a glassy pellet and an overlying high-density gradient at the bottom of the tube. The vesicles of interest loosely congregate above the glassy pellet with the residual Percoll. Concentrating the fraction by centrifuging onto a high-concentration sucrose cushion can help because Percoll will sediment through the cushion. Alternatively, the vesicles of interest can be concentrated by flotation

BGU ILLEGRA UNIVERSITY LIBRARY

from a medium of higher buoyant density in which Percoll pellets. Finally, for SDS-PAGE of samples that contain residual Percoll, heating 5 to 10 min in sample buffer and then microcentrifuging 10 min at top speed prior to loading on the gel will remove nearly all remaining Percoll.

Gradient fractionation using D_2O

A main concern in using D_2O for fractionation is maintaining osmotic balance across organelle membranes so that lysis or leakage of contents is avoided. Gumbiner and Kelly (1981) include sucrose in their density gradient solutions to avoid lysis of secretion granules, whereas Wagner et al. (1978) include NaCl to stabilize synaptic vesicles. A similar strategy will be needed if fractionation in D_2O (or glycerol which also permeates membranes) is applied to other organelles.

Gradient fractionation using nonionic iodinated solutes

Solutions of metrizamide and Nycodenz should be stored in the dark for stability. As with other density gradient procedures, the concentrations of solutions used to prepare gradients should be standardized using a refractometer. In addition, the densities of gradient fractions collected after centrifugation can be determined by measuring refractive index.

Gel filtration

The column resin needs to be pretreated to suppress loss of small vesicles due to adsorption. This problem is particularly acute the first few times the column is used. In using controlled-pore glass (CPG-3000), Carlson and Kelly (1978) recommend treating the column periodically with polyethylene glycol (PEG; Carbowax) to suppress adsorption. The principle is the same as that used commercially in coating colloidal silica particles with polyvinylpyrrolidone in preparing Percoll. Also, pretreating the resin (either Sephacryl or CPG) by running a sample containing small phospholipid vesicles helps to suppress subsequent adsorptive losses.

Satisfactory and reproducible gel filtration of vesicles requires thorough dispersion of the starting sample. Most often this is a microsomal (small vesicle) pellet. After resuspending the pellet, it should be thoroughly dispersed by Dounce homogenization or by passage several times though a small-bore hypodermic needle (>25G). Briefly microcentrifuge the dispersed sample to pellet any residual aggregates and load the supernatant.

Preparative and analytical electrophoretic separation

To ensure that aggregation of starting material is minimal, after thorough dispersal of microsomal pellets, the samples should be microcentrifuged briefly before loading.

Setting the flow rate of the peristaltic pump used to drain the collecting well of the preparative agarose electrophoresis device for efficient clearance of separated organelles is also essential.

Solid-phase immunoadsorption

The antibody immobilized on the solid phase support must be both stably bound and active. Generally, if an antibody is good for immunoprecipitation, it is a candidate for use in immunoadsorption (contingent on epitope availability). Mouse monoclonal antibodies that have a low affinity for protein A can be adapted for immunoadsorption by using either beads coated with secondary (anti-mouse antibody) or immobilized protein A and a rabbit anti-mouse antibody bridge.

Always run a parallel control in which an irrelevant antibody is substituted for the specific antibody. This will indicate the level of nonspecific adsorption in the procedure.

Before beginning the immunoadsorption, removal of any organelle aggregates is essential. This is especially critical where centrifugation is used to collect and wash the solid phase. As in the other procedures, thorough dispersal followed by brief centrifugation can be used to eliminate aggregates (which are hopefully only a minor fraction of antigen).

For large and/or dense organelles (e.g., stacked Golgi complexes and exocrine secretion granules), it is advisable to use a plate magnet to immobilize the magnetic beads and adsorbed organelles during washing. Repeated packing of the beads by centrifugation or in the commercially available magnetic concentrator used with microcentrifuge tubes can dislodge these organelles. Furthermore, large organelles may sediment independent of their specific association with the immunoadsorbant.

Purification by lectin adsorption

Identification of a lectin that will bind to the plasma membrane of the cells of interest is essential to the procedure. Efficient elution of

PCU IASE CENTRAL UNIVERSITY LIBRARY

bound vesicles is also important; N-acetylglucosamine may not effectively compete for multiple lectin interaction sites per vesicle. An assay for a plasma membrane marker is useful to judge recovery, i.e., the amounts of membrane bound, not bound, and eluted by the competing ligand with respect to the original amount used for adsorption.

It is advisable to check marker activities for organelles that are not expected to adhere to the lectin to evaluate potential nonspecific adsorption of contaminating organelles.

Density shift using digitonin

As digitonin at higher concentrations behaves like a detergent and lyses most membrane-bounded organelles, a preliminary titration experiment is necessary to identify sublytic digitonin concentrations for use in density shift (see Basic Protocol 6, step 4).

Density shifts may be fairly subtle, especially where the cholesterol content of the membrane is low. Thus it is important to assay carefully the complete distribution of all markers and the protein (or other macromolecule) of interest across the density gradient.

Density shift using colloidal gold conjugates

In preparing colloidal gold conjugates, it is essential to identify stabilizing conditions in which the protein coating on the gold particles is sufficient to prevent flocculation. A simple test is described in the protocol.

As in most other procedures, minimal aggregation among organelles following homogenization is essential, otherwise, density perturbation extends to organelles that are associated artifactually with the particular density-modified compartment.

Extraction of extrinsic proteins from membranes using sodium carbonate

Sodium carbonate treatment is most effective if the content of other electrolytes in the sample is low. This is the reason membrane pellets are resuspended in sucrose medium with only a low buffer content and then diluted with a large volume of the high-pH carbonate solution. Carbonate induces discontinuities in the membrane bilayer (see Fujiki et al., 1982) that may be important to effectively removing extrinsic proteins. Higher concentrations of NaCl (and possibly other electrolytes) reduce the formation of discontinuities. Buffers of course decrease the pH elevation.

Do not neglect to include the sucrose cushion through which the membranes are spun after carbonate treatment. Membranes pelleted directly in carbonate are rock-like and extremely difficult to resuspend.

Anticipated Results

The key parameters to consider in judging the success of various preparative fractionation procedures are the yield of material and extent of purification. Yield is generally assessed in two ways. First, assays of organelle marker activities are used to determine what fraction of the total activity present in the original tissue homogenate is recovered in the purified fraction. This is a measure of whether the fraction is representative; generally recoveries >10% of total in the fraction of interest are satisfactory, especially as recovery is usually sacrified in the interest of purity. A complementary assessment in definitive studies involves analysis of the fractional contamination by undesirable organelles. The second measure is more practical and involves determining the amount of protein recovered in the fraction from the amount of tissue used. The extent of purification is usually judged by evaluating the enrichment of a marker activity from the homogenate to the final fraction. It is usually expressed as a ratio of activity per mg protein in the final fraction/activity per mg protein in the homogenate. For organelles that constitute large fractions of total protein—e.g., 20% for ER and mitochondria in liver (Wattiaux et al., 1978) and >25% for secretion granules in exocrine cells (Zastrow and Castle, 1987)—enrichment cannot exceed a factor of 4 or 5. However, if an organelle constitutes only 1% to 2% of total cell protein (e.g., Golgi and lysosomes in liver), then theoretical enrichments of 50- to 100-fold are possible.

For analytical procedures, the success of separations is mainly judged by the comparative distributions of marker activities for distinct organelles. Examples of anticipated results are given in Figure 4.2.1 (rate zonal centrifugation), Figure 4.2.6 (analytical electrophoresis on a density gradient), and Figure 4.2.7 (density shift using digitonin).

Examples of expected results, particularly for the preparative centrifugation techniques, are given in the following discussion. Good yields but low purity should be expected when using differential centrifugation by velocity (Basic Protocol 1). The microsomal fraction constitutes 13% of homogenate protein (18 to 20 mg protein per gram of liver used) and contains 40% to 50% of ER marker activities (glucose-6-phosphatase and ER-associated es-

BC INSTITUTE CENTRAL UNIVERSITY LIBRARY

terase), 40% of a Golgi marker (galactosyl transferase), and 25% of a plasma membrane marker activity (5′-nucleotidase; Bergeron et al., 1982). By using differential centrifugation by equilibrium (Basic Protocol 2) subsequent to differential centrifugation by velocity, rough microsomes (derived mainly from rough ER) can be efficiently purified with a yield of 7.5 to 10 mg protein per gram of liver used (Adelman et al., 1973; Bergeron et al., 1982). Mitochondria can be purified by repeated differential centrifugation by velocity (Basic Protocol 1) with a yield of 15 mg protein per gram of rat liver used (Carvalho-Guerra, 1974). Unfortunately the purity was not reported, but it would not be expected to exceed 5-fold for reasons discussed above.

Impressive purifications but often lower yields are expected using differential centrifugation by equilibrium (Basic Protocol 2) or its variations (e.g., Alternate Protocol 4). The intact Golgi fraction used as an example in Basic Protocol 2 is more than 100-fold purified from the homogenate (close to the theoretical limit, as Golgi represents close to 1% of liver protein (Wattiaux et al., 1978)). The yield of total galactosyl transferase is unusually high (33% of total); however, the absolute amount of material obtained is low (0.24 mg protein per gram liver), as expected (Bergeron et al., 1982). In isolating hepatic lysosomes using gradient centrifugation in iodinated nonionic solutes (Alternate Protocol 4), Wattiaux et al. (1978) obtained a 10% yield of lysosomal marker activities in their purest fraction and the absolute amount of protein was ~0.25 mg per gram of liver used. However, the purity of the fraction (65- to 80-fold) is extremely impressive, as lysosomes are estimated to be 1% to 1.4% of total liver protein. The procedure of Wattiaux et al. (1978) is a model for careful accounting of recovery of activities and analysis of cross-contamination of fractions by other organelles.

Yields using Percoll gradient centrifugation (Alternate Protocol 2) generally can be expected to be quite good. For the secretion granules used as an example, about 30% of the secretory protein amylase is recovered in the purified fraction. Generally 20% to 25% of amylase is released from granules that lyse during homogenization, while another 20% fails to be released from cells during homogenization; the latter sediments in the nuclear pellet (see Zastrow and Castle, 1987, and references therein). For secretion granules, which are much more dense than other cytoplasmic organelles, the purity of fractions obtained using Percoll gradients can be expected to be quite good, provided conditions of centrifugation (Percoll density, times of runs) are used reproducibly. Where organelles of interest and contaminants have more similar densities, the possibility of contamination can be significant and needs to be considered. Even though the forte of Percoll is generating very shallow density gradients with good spatial resolution, biological organelles are heterogeneous (either intrinsically or especially after cell homogenization) and distribution of a particular organelle over a range of densities can be expected.

Gel filtration to isolate secretory vesicles (Basic Protocol 3) is a procedure that is restricted to small organelles. Losses due to vesicle adsorption to the column matrix can be expected to be significant (possibly as much as 50%) during initial use of the column but should improve significantly thereafter. Even though vesicle purification using the column can be impressive, homogeneity should not be expected. Homogenization generates small vesicles from larger organelles and these will co-fractionate with the secretory (or synaptic) vesicles. Generally it is wise to follow the distribution of marker activities in the fractions eluted from the column (as in Fig. 4.2.3) if co-localization is being evaluated.

The use of electrophoretic techniques (e.g., Basic Protocol 4 and Alternate Protocol 5), although quite promising, is still rather limited. The initial efforts suggest that both recovery and purity using the preparative procedure may be quite high for selected organelles (Kedersha and Rome, 1986). However, more widespread use is needed before it is possible to establish expectations for yield and purity. Lectin-based purification of plasma membrane fractions (Alternate Protocol 6) also has not been employed both widely and quantitatively. For the exocrine pancreatic apical plasma membrane fraction used as an example, yields were not reported. However, the fact that most plasma membrane vesicles are oriented with ectodomain outward (and thus accessible to the lectin) suggests that the yield could potentially be good. The purification (75-fold enrichment over the homogenate) is very good (Paul et al., 1992).

Density-shift using colloidal gold conjugates (Basic Protocol 7) appears to be quite promising for obtaining fractions enriched in different segments of the endocytic pathway in both quite good yield and purity (Beaumelle et al., 1990). The yields reported for plasma membranes, endosomes, coated pits, and lysosomes

isolated by density modification are 40%, 32%, 10%, and 16%, respectively. Contamination by organelles outside the endocytic pathway appears negligible based on marker assays. Cross-contamination among endocytic compartments also appears to be low but potentially is limited by the extent to which individual compartments can be density-modified selectively in the cells prior to fractionation.

Extraction of extrinsic proteins from membranes using sodium carbonate (Basic Protocol 8) generally can be expected to be 90%, while retention of integral proteins in sedimentable membranes is usually quantitative. A result suggesting partial extraction will be difficult to interpret. If established integral membrane proteins are underrecovered the final membrane pellets, then incomplete recovery of membranes following centrifugation needs to be considered. On the other hand, if the loss of the proteins of interest is selective with respect to the markers, then these proteins may be extrinsic but inefficiently extracted. In this case, distribution should be judged following a repeat extraction.

Time Considerations

This section provides estimates of the time required for organelle purification by the procedures that are likely to be used most frequently. For the other protocols, total times and convenient breakpoints can be easily estimated from the steps that are listed.

In differential centrifugation by velocity (Basic Protocol 1), sample preparation and homogenization takes about 20 min; centrifugation to obtain the PNS takes 10 min; ultracentrifugation to obtain the crude mitochondrial fraction requires 30 min, and further purification takes another 30 min; finally, ultracentrifugation to obtain the microsomal pellet requires 2 hr. For each step requiring resuspension of pellets, include an additional 10 min.

In differential centrifugation by equilibrium (Basic Protocol 2), preparation of the density gradients requires 30 to 60 min in advance of starting the fractionation. Sample preparation and homogenization takes about 20 min while loading the gradients takes about 15 to 30 min depending on experience. Running the gradient takes 3 hr (with an additional 15 to 30 min to allow the rotor to coast to a stop) in the example given; however, in most equilibrium runs this step goes overnight. Allow ~30 min to collect the bands from the gradients after the run, 90 min to pellet them by additional centrifugation, and 15 min to resuspend them for use or storage. In many procedures, the gradient step is used in series with differential centrifugation by velocity, so selected times as in Basic Protocol 1 need to be added in.

In rate zonal centrifugation (Alternate Protocol 1), preparation of the gradients requires 30 to 60 min in advance of starting the fractionation. Sample preparation and homogenization takes 20 to 30 min (more time is allotted for surgery to remove pancreas, which is more difficult than for liver); low-speed centrifugations to prepare the PNS require 30 min together; loading gradients takes 15 to 30 min; gradient centrifugation requires 2 hr plus a 30-min coast of the rotor to stop; and collection of the fractionations either manually or with a gradient-elution device requires 10 to 15 min per centrifuge tube. Any additional centrifugation to pellet fractions generally requires 1 to 1.5 hr, with 15 min to resuspend pellets at the end.

For gradient fractionation using Percoll (Alternate Protocol 2), allow 60 min for sample preparation, homogenization, and generation of PNS as above. Any adjustment of density requiring measurement of refractive index takes ~10 min, while centrifugation of the gradients requires 30 min to 2 hr, depending on the organelle being isolated. Collection of fractions after the run requires 5 or 15 min per centrifuge tube for manual collection or collection with a gradient-elution device, respectively. Subsequent centrifugation to separate most of the Percoll from the fractions requires 45 min in a low-speed centrifuge or 1 to 2 hr if sedimentation in an ultracentrifuge is used.

Gradient fractionation using nonionic iodinated solutes (Alternate Protocol 4) requires similar amounts of time as above for sample preparation, homogenization, and centrifugation to generate a PNS. In the example protocol given, the rehomogenization and recentrifugation of the initial nuclear pellet adds another 20 min to give a total time of ~60 min to prepare the PNS for loading on gradients. A 10,000-rpm centrifugation to pellet mitochondria requires an additional 30 min. Continuous gradients can be made in advance of starting the experiment whereas discontinuous gradients are usually made right before use during the low-speed centrifugation steps; allow 10 min per tube. Ultracentrifugation of the gradients requires 2.5 to 3 hr (including coast of the rotor to a stop). Unloading gradients takes 10 to 15 min per tube. Any additional centrifugation to pellet

BOURNEMOUTH UNIVERSITY LIBRARY

the isolated fractions requires 1.5 hr plus 5 min per tube to resuspend the final pellets for assays or storage.

For gel filtration (Basic Protocol 3), sample preparation and centrifugation to prepare the column load approximately follows the time-course for differential centrifugation by velocity. Where yeast are used as a source of sample, allow 60 min to spheroplast and harvest the yeast by centrifugation before starting the homogenization. The gel-filtration column is run overnight; fractions are maintained at 4°C during assay, and then an additional 1 to 1.5 hr is needed for pooling the selected fractions and pelleting them by centrifugation.

For electrophoretic separations (Basic Protocol 4 and Alternate Protocol 5), about 3 to 4 hr is required for preparing and homogenizing tissue and sedimenting the microsomal starting material for the electrophoretic separation. Centrifugation times are similar to those in Basic Protocol 1, and additional time is included to allow for assay of total protein and brief trypsinization (Alternate Protocol 5). The agarose gel for preparative electrophoresis (Basic Protocol 4) can be prepared in advance of the experiment and requires at least 2 hr to solidify before use. The density gradient for analytical electrophoresis (Alternate Protocol 5) requires ~30 min and is set up just before loading. Preparative electrophoresis in agarose is conducted for 16 hr, while analytical electrophoresis on the density gradient will take an hour, including the solution changes during the run. Fractions for assay and further processing are collected continuously during electrophoresis on the agarose gel whereas fractions are collected at the end of the run on the analytical density gradient, requiring 30 min.

Solid-phase immunoadsorption using antibody-coated magnetic beads (Basic Protocol 5) requires 4 hr overnight incubation in advance of the experiment to coat the beads with antibody and subsequently wash them. During the experiment, preparation of the PNS requires about 30 min (as in Basic Protocol 1). Subsequent centrifugation to prepare membranes to load on the gradient requires 1.5 hr (including loading and unloading the tubes) in the example given. Preparation and loading of discontinuous gradients, centrifugation, and collection of fractions requires a total of 2 to 2.5 hr. Immunoadsorption of organelles requires 3 hr and subsequent washing needs another 30 min before samples are harvested or used in functional assays.

Extraction of extrinsic proteins from membranes using sodium carbonate (Basic Protocol 8) can be performed either immediately after preparation by any of the protocols listed or after storage of samples that have been resuspended in buffered sucrose and frozen. Carbonate treatment requires 30 min; loading of treated samples for pelleting through a sucrose cushion requires 5 min per centrifuge tube; and centrifugation requires 1 to 2 hr. Collecting the final supernatants and resuspended pellets requires 5 min per tube.

Literature Cited

Adelman, M.R., Blobel, G., and Sabatini, D.D. 1973. An improved cell fractionation procedure for the preparation of rat liver membrane-bound ribosomes. *J. Cell Biol.* 56:191-205.

Ajioka, R.S. and Kaplan, J. 1987. Characterization of endocytic compartments using the horseradish peroxidase/diaminobenzidene density shift technique. *J. Cell Biol.* 104:77-85.

Amar-Costasec, A., Wibo, M., Thines-Sempoux, Beaufay, H., and Berthet, J. 1974. Analytical study of microsomes and isolated subcellular membranes from rat liver. *J. Cell Biol.* 62:717-745.

Amigorena, S., Drake, J.R., Webster, P., and Mellman, I. 1994. Transient accumulation of new class II MHC molecules in a novel endocytic compartment in B lymphocytes. *Nature* 369:113-120.

Balch, W.E. and Rothman, J.E. 1985. Characterization of protein transport between successive compartments of the Golgi apparatus: Asymmetric properties of donor and acceptor activities in a cell-free system. *Arch. Biochem. Biophys.* 240:413-425.

Balch, W.E., Dunphy, W.G., Braell, N.A., and Rothman, J.E. 1984. Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of *N*-acetylglucosamine. *Cell* 39:405-416.

Barlowe, C., Orci. L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., Rexach, M.F., Ravazzola, M., Amherdt, M., and Schekman, R. 1994. Cop II: A membrane coat formed by sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77:895-907.

Baykov, A.A., Evtushenko, O.A., and Avoeva, S.M. 1988. A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase–based enzyme immunoassay. *Anal. Biochem.* 171:266-270.

Beardmore, J., Howell, K.E., Miller, K., and Hopkins, C.R. 1987. Isolation of an endocytic compartment from A431 cells using a density modification procedure employing a receptor-specific monoclonal antibody complexed with colloidal gold. *J. Cell Sci.* 87:495-506.

CU IASTATE LIBRARIES

Beaufay, H., Bendall, D.S., Baudhuin, P., Wattiaux, R., and DeDuve, C. 1959. Tissue fractionation studies: 13. Analysis of mitochondrial fractions from rat liver by density gradient centrifuging. *Biochem. J.* 73:628-637.

Beaumelle, B.D., Gibson, A., and Hopkins, C.R. 1990. Isolation and preliminary characterization of the major membrane boundaries of the endocytic pathway in lymphocytes. *J. Cell Biol.* 111:1811-1823.

Becker, D. M. and Lundblad, V. 1994. Introduction of DNA into yeast cells. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, eds.) pp. 13.7.1-13.7.10. John Wiley and Sons, New York.

Bergeron, J.J.M., Rachubinski, R.A., Sikstrom, R.A., Posner, B.I., and Paiement, J. 1982. Galactose transfer to endogenous acceptors within Golgi fractions of rat liver. *J. Cell Biol.* 92:139-146.

Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256:1604-1607.

Carlson, S., Wagner, J.A., and Kelly, R.B. 1978. Purification of synaptic vesicles from elasmobranch electric organ and the use of biophysical criteria to demonstrate purity. *Biochemistry* 17:1188-1199.

Carvalho-Guerra, F. 1974. Rapid isolation techniques for mitochondria: Technique for rat liver mitochondria. *Methods Enzymol.* 31:299-305.

Courtoy, P.J., Quintart, J., and Baudhuin, P. 1984. Shift of equilibrium density induced by 3,3′-diaminobenzidene cytochemistry: A new procedure for the analysis and purification of peroxidase-containing organelles. *J. Cell Biol.* 98:870-876.

DeDuve, C. and Berthet, J. 1954. The use of differential centrifugation in the study of tissue enzymes. *Int. Rev. Cytol.* 3:225-275.

Emmelot, P. and Bos, C.J. 1966. Studies on plasma membranes. II. K+-dependent *p*-nitrophenyl phosphatase activity of plasma membranes isolated from rat liver. *Biochim. Biophys. Acta* 121:375-385.

Fujiki, Y., Hubbard, A.L., Fowler, S., and Lazarow, P.B. 1982. Isolation of intracellular membranes by means of sodium carbonate treatment: Application to endoplasmic reticulum. *J. Cell Biol.* 93:97-102.

Futter, C. and Hopkins, C.R. 1989. Subfractionation of the endocytic pathway: Isolation of compartments involved in the processing of internalized epidermal growth factor-receptor complexes. *J. Cell Sci.* 94:685-694.

Futter, C.E., Connolly, C.N., Cutler, D.E., and Hopkins, C.R. 1995. Newly synthesized transferrin receptors can be detected in the endosome before they appear on the cell surface. *J. Biol. Chem.* 270:10999-11003.

Gruenberg, J.E. and Howell, K.E. 1986. Reconstitution of vesicle fusions occurring in endocytosis with a cell-free system. *EMBO J.* 5:3091-3101.

Gumbiner, B. and Kelly, R.B. 1981. Secretory granules of an anterior pituitary cell line, AtT-20, contain only mature forms of corticotropin and β-lipotropin. *Proc. Natl. Acad. Sci. U.S.A.* 78:318-322.

Gupta, D. and Tartakoff, A.M. 1989. Lectin–colloidal gold–induced density perturbation of membranes: Application to affinity elimination of the plasma membrane. *Methods Cell Biol.* 31:247-263.

Handley, D.A., Arbeeny, C.M., Witte, L.D., and Chen, S. 1981. Colloidal gold–low density lipoprotein conjugates as membrane receptor probes. *Proc. Natl. Acad. Sci. U.S.A.* 78:368-371.

Havel, R.J., Eder, H.A., and Bragdon, J.H. 1955. The distribution and chemical composition of ultracentrifugally separated lipoproteins in human serum. *J. Clin. Invest.* 34:1345-1353.

Helmy, S., Porter-Jordan, K., Dawidowicz, E.A., Pilch, P., Schwartz, A.L., and Fine, R.E. 1986. Separation of endocytic from exocytic coated vesicles using a novel cholinesterase-mediated density shift technique. *Cell* 44:497-506.

Holcomb, C., Etcheverry, T., and Schekman, R. 1987. Isolation of secretory vesicles from *Saccharomyces cerevisiae*. *Anal. Biochem.* 166:328-334.

Howell, K.E., Schmid, R., Ugelstad, J., and Gruenberg, J. 1989. Immunoisolation using magnetic solid supports: Subcellular fractionation for cell-free functional studies. *Methods Cell Biol.* 31:265-292.

Huttner, W.B., Schlieber, W., Greengard, P., and DeCamilli, P. 1983. Synapsin I (Protein I), a nerve terminal-specific phosphoprotein: III. Its association with synaptic vesicles studied in a highly purified synaptic vesicle preparation. *J. Cell Biol.* 96:1374-1388.

Kedersha, N.L. and Rome, L.H. 1986. Preparative agarose gel electrophoresis for the purification of small organelles and particles. *Anal. Biochem.* 156:161-170.

Laurie, S.M., Cain, C.C., Lienhard, G.E., and Castle, J.D. 1993. The glucose transporter GluT4 and secretory carrier membrane proteins (SCAMPs) colocalize in rat adipocytes and partially segregate during insulin stimulation. *J. Biol. Chem.* 268:19110-19117.

Leelavathi, D.E., Estes, L.W., Feingold, D.S., and Lombardi, B. 1970. Isolation of a Golgi-rich fraction from rat liver. *Biochim. Biophys. Acta* 211:124-138.

Lesley, J., Domingo, D.L., Schulte, R., and Trowbridge, I.S. 1984. Effect of an anti-murine transferrin receptor–ricin A conjugate on bone marrow stem and progenitor cells treated in vitro. *Exp. Cell Res.* 150:400-407.

Loh, Y.P., Tam, W.W.H., and Russell, J.T. 1984. Measurement of pH and membrane potential in secretory vesicles isolated from bovine pituitary intermediate lobe. *J. Biol. Chem.* 259:8238-8245.

PENNSYLVANIA STATE UNIVERSITY LIBRARY

Marsh, M., Schmid, S., Kern, H., Harms, E., Male, P., Mellman, I., and Helenius, A. 1987. Rapid analytical and preparative isolation of functional endosomes by free flow electrophoresis. *J. Cell Biol.* 104:875-886.

Merisko, E., Farquhar, M.G., and Palade, G.E. 1981. Coated vesicle isolation by immunoadsorption on *Staphylococcus aureus* cells. *J. Cell Biol.* 92:846-857.

Mueller, S.C. and Branton, D. 1984. Identification of coated vesicles in *Saccharomyces cerevisiae*. *J. Cell Biol.* 98:341-346.

Orlowsky, M. and Meister, A. 1963. γ-glutamyl-*p*-nitroanilide: A new convenient substrate for the determination and study of L- and D-γ-glutamyltranspeptidase activities. *Biochim. Biophys. Acta* 73:679-681.

Paul, E., Hurtubise, Y., and LeBel, D. 1992. Purification and characterization of the apical plasma membrane of the rat pancreatic acinar cell. *J. Memb. Biol.* 127:129-137.

Rickwood, D. 1984. Centrifugation, a Practical Approach (2nd ed.). IRL Press, Oxford.

Roep, B.O., Arden, S.D., deVries, R.R.P., and Hutton, J.C. 1990. T-cell clones from a type 1 diabetes patient respond to insulin secretory granule proteins. *Nature* 345:632-634.

Rubenstein, J.L.R., Fine, R.E., Luskey, B.D., and Rothman, J.E. 1981. Purification of coated vesicles by agarose gel electrophoresis. *J. Cell Biol.* 89:357-361.

Salamero, J., Sztul, E.S., and Howell, K.E. 1990. Exocytic transport vesicles generated in vitro from the trans-Golgi network carry secretory and plasma membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* 87:7717-7721.

Saucan, L. and Palade, G.E. 1994. Membrane and secretory proteins are transported from the Golgi complex to the sinusoidal plasmalemma of hepatocytes by distinct vesicular carriers. *J. Cell Biol.* 125:733-741.

Slot, J.W. and Geuze, H.J. 1985. A new method of preparing gold probes for multiple-labeling cytochemistry. *Eur. J. Cell Biol.* 38:87-93.

Tulp, A., Verwoerd, D., and Pieters, J. 1993. Application of an improved density gradient electrophoresis apparatus to the separation of proteins, cells and subcellular organelles. *Electrophoresis* 14:1295-1301.

Tulp, A., Verwoerd, D., Dobberstein, B., Ploegh, H.L., and Pieters, J. 1994. Isolation and characterization of the intracellular MHC class II compartment. *Nature* 369:120-126.

Wagner, J.A., Carlson, S.S., and Kelly, R.B. 1978. Chemical and physical characterization of cholinergic synaptic vesicles. *Biochem.* 17:1199-1206.

Walworth, N.C. and Novick, P.J. 1987. Purification and characterization of constitutive secretory vesicles from yeast. *J. Cell Biol.* 105:163-174.

Warnock, D.E., Roberts, C., Lutz, M.S., Blackburn, W.A., Young, W.W., and Baenziger, J.U. 1993. Determination of plasma membrane lipid mass and composition in cultured Chinese hamster ovary cells using high gradient magnetic affinity chromatography. *J. Biol. Chem.* 268:10145-10153.

Wattiaux, R., Wattiaux-DeConnick, S., Ronveaux-Dupal, M.-F., and Dubois, F. 1978. Isolation of rat liver lysosomes by isopycnic centrifugation in a metrizamide gradient. *J. Cell Biol.* 78:349-368.

Zastrow, M.V. and Castle, J.D. 1987. Protein sorting among two distinct export pathways occurs from the content of maturing exocrine storage granules. *J. Cell Biol.* 105:2675-2684.

Contributed by J. David Castle
University of Virginia
Charlottesville, Virginia

UCU IASI / CENTRAL UNIVERSITY LIBRARY

Subcellular Fractionation of Tissue Culture Cells

The development of cell fractionation techniques over the last few decades has provided the means to analyze the composition and properties of purified cellular elements. In particular, subcellular fractionation is essential for the development of cell-free assays that reconstitute complicated cellular processes. These assays have provided new and important tools to understand the molecular mechanisms of complex cellular functions permitting these functions to be studied in the test tube as a series of biochemical reactions.

General methods for fractionating mammalian tissues are presented in *UNIT 4.2*. The protocols in this unit describe fractionation of tissue culture cells to immunoisolate early and late endosomes under conditions where these compartments retain their capacity to support membrane transport in vitro. To prepare different populations of endosomes, monolayers of cells are exposed to a fluid-phase marker for endosomal compartments (horseradish peroxidase, HRPO) either briefly to label early endosomes or with a chase period to label endosomal carrier vesicles or late endosomes (see Basic Protocol 1). Alternatively, vesicular stomatitis virus (VSV) can be fused with the cells to add VSV G protein to the cell membranes (see Basic Protocol 2) and the cells then exposed to HRPO. Labeled cells from either protocol are then homogenized (see Basic Protocol 3) and the postnuclear supernatant (PNS) is fractionated by a flotation gradient (see Basic Protocol 4). Specific endosomal compartments are isolated using magnetic or polyacrylamide beads coupled to an antibody specific for VSV G protein (see Basic Protocol 5). The Support Protocol describes growth conditions for preparing BHK-21 cells for isolation of endosomes. The Commentary discusses preparation of a "balance sheet" for the isolation procedure (see Background Information) and problems that may be encountered in subcellular fractionation of tissue culture cells (see Critical Parameters as well as the Commentary in *UNIT 4.2*).

INTERNALIZATION OF FLUID-PHASE MARKER

BASIC PROTOCOL 1

Several proteins, including rab GTPases and receptors (see Table 4.3.1), can be used to follow early or late endosomes. In the absence of appropriate reagents, markers endocytosed in the fluid phase can allow convenient discrimination between the compartments (see Table 4.3.1). Horseradish peroxidase (HRPO) is widely used as a marker for fluid-phase endocytosis. HRPO internalized at a concentration of 0.5 to 10 mg/ml distributes uniformly within the lumen of endosomal compartments and is easily measured using a colorimetric reaction. Internalization is carried out directly in the tissue culture dish for cells growing attached to the substratum. For suspension cells, the method is essentially identical, except that all changes of buffers and solutions must be accomplished by centrifugation. Once labeled, the cells must be washed thoroughly to remove nonspecifically adsorbed marker. Washed cells may be incubated further in the absence of marker to label endosomal carrier vesicles and late endosomes.

Materials

Monolayer cultures of BHK-21 cells in 10-cm tissue culture plates, 14 to 18 hr after plating (see Support Protocol 3)
PBS (*APPENDIX 2E*), ice cold
2 to 10 mg/ml horseradish peroxidase (HRPO) in internalization medium (see recipe), 37°C
PBS/BSA: PBS containing 5 mg/ml BSA, ice cold and 37°C
IM/BSA: Internalization medium (see recipe) containing 2 mg/ml BSA, 37°C

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON COLLEGE CENTRAL UNIVERSITY LIBRARY

Table 4.3.1 Labeling Conditions for Characterization of the Different Endosomal Compartments[a,b]

Endosomal compartment	Duration of incubation at 37°C	Markers	References
EE	5 min	Rab5, rab4	Chavrier et al. (1990); Gorvel et al. (1991)
		Annexin II	Emans et al. (1993)
		Receptors recycling with the plasma membrane (e.g., transferrin-R)	Gruenberg and Howell (1989); Emans et al. (1991)
ECV	5-min pulse + 45-min chase after depolymerization of microtubules		Gruenberg et al. (1989); Bomsel et al. (1990); Aniento et al. (1993)
LE	5-min pulse + 45-min chase	Rab7, rab9	Chavrier et al. (1990); Gorvel et al. (1991)
		CI-MPR	Griffiths et al. (1988); Aniento et al. (1993a)
		lgps	Geuze et al. (1983); Griffiths et al. (1988); Gorvel et al. (1991)

[a]Several proteins listed here exhibit a restricted distribution within endosomes, but can also be found in other compartments (e.g., MPR in TGN and late endosomes; rab5 in plasma membrane and early endosomes). Analysis of purified fractions must thus combine different markers to identify endosomes unambiguously.

[b]Abbreviations: transferrin-R transferrin receptor; CI-MPR, cation-independent mannose-6-phosphate receptor; lgps, lysosomal glycoproteins; TGN, trans Golgi network. The 4a1 and 2a5 monoclonal antibodies recognize two lgps that are abundant in late endosomes (Aniento et al., 1993).

Cold plate: wet metal plate lying flat on top of ice in an ice bucket
Rocking platform
Warm plate: 37 °C water bath with a flat metal plate slightly below the water level
or 37°C incubator

1. Place tissue culture dishes containing monolayer cultures of BHK-21 cells on the cold plate. Aspirate the medium and add 5 to 10 ml ice-cold PBS to each dish. Incubate 2 to 3 min on a rocking platform. Repeat this step twice to thoroughly wash the cells.

   Use two to three dishes of cells per time point. Keep the dishes level on the plate in the ice bucket.

2. Aspirate the PBS and prewarm each dish 1 to 2 sec on the warm plate. Add 2.5 ml prewarmed HRPO to the dish and incubate 5 min on the warm plate.

   In most cases, HRPO at a concentration of 2 mg/ml provides sufficient labeling. If necessary, higher concentrations (up to 10 mg/ml) can be used to increase the signal (for example, when fractions are further purified by immunoisolation).

   To reduce costs, small volumes of HRPO (e.g., 2.5 ml) can be used for incubation at 37°C. The culture dish should be placed on a metal plate mounted on a rocking platform in a 37°C incubator to ensure that the cells are adequately covered with liquid during the incubation.

   After a 5-min incubation at 37°C, membrane proteins and fluid-phase tracers distribute predominantly within early endosomes (Gruenberg et al., 1989; Gorvel et al., 1991; Emans et al., 1993). After longer internalization times, the markers appear first in endosomal carrier vesicles and then sequentially in late endosomes and lysosomes (Bomsel et al., 1990; Aniento et al., 1993a).

3. Return the dishes to the cold plate. Aspirate the HRPO and wash each dish three times, 5 to 10 min each, with 5 ml PBS/BSA. Aspirate the PBS/BSA and wash each dish

PCU IASI / CENTRAL UNIVERSITY LIBRARY

twice with 5 ml PBS. To label endosomal carrier vesicles or late endosomes, proceed to steps 4 and 5. Otherwise, proceed to homogenization (see Basic Protocol 3).

4. *Optional:* If the experiment requires fluid-phase labeling of endosomal carrier vesicles or late endosomes, immediately after the internalization step rinse each dish two or three times with 5 ml of 37°C PBS/BSA. Add 10 ml IM/BSA and incubate 45 min at 37°C.

   These pulse-chase conditions are preferred to continuous internalization of HRPO because under these conditions the amount of HRPO in earlier compartments is reduced.

5. Return the dishes to the cold plate. Then, as in step 3 above, aspirate the medium and wash each dish twice, 5 min each, with 5 ml PBS/BSA. Aspirate the PBS/BSA and rinse twice with PBS.

   Markers can also be allowed to accumulate in the intermediate endosomal carrier vesicles by preincubating the cells 1 to 2 hr at 37°C in medium containing 10 μM nocodazole to depolymerize microtubules. The conditions for fluid-phase labeling and marker internalization are the same, except that all wash solutions and media used, including internalization and chase media, should contain 10 μM nocodazole. Alternatively, endosomal carrier vesicles can be labeled by incubating the cells 2 hr at 20°C without nocodazole. However, internalization rates and labeling efficiency will be higher at 37°C than at 20°C.

IMPLANTATION AND INTERNALIZATION OF VSV G PROTEIN

BASIC PROTOCOL 2

The transmembrane glycoprotein G of vesicular stomatitis virus (VSV), or another transmembrane endosomal antigen for which there is an antibody, can be used to immunoisolate endosomal fractions (see Fig. 4.3.1). Virus is first bound to the plasma membrane by wheat germ agglutinin–mediated adhesion, then fused with the plasma membrane by a brief exposure to low pH medium that inhibits endocytosis (Davoust et al., 1987). The cells are then washed and incubated with HRPO at 37°C. Under these conditions, the G protein, a membrane marker, is co-internalized with HRPO, a fluid-phase marker. The G protein cytoplasmic domain becomes exposed on the surface of endosomal elements and is accessible to antibodies after homogenization.

Materials

Monolayer cultures of BHK-21 cells in 10-cm tissue culture plates, 14 to 18 hr after plating (see Support Protocol 3)
100 μg/ml wheat germ agglutinin (WGA) from *Triticum vulgaris* (Sigma) in 7.4 M medium (see recipe), ice cold
20 μg/ml vesicular stomatitis virus (VSV; see Support Protocol 1) in 7.4 M medium (see recipe), ice cold
Fusion medium: MEM containing 20 mM succinate, pH 4.9, 37°C
60 mM *N*-acetylglucosamine (GlcNAc) in 7.4 M medium (see recipe), ice cold
Vacuum aspirator connected to a trap containing disinfectant (e.g., chloramine T)
Additional reagents and equipment for fluid-phase internalization (see Basic Protocol 1)

CAUTION: VSV is a rodent pathogen. Consult with the local hazardous materials authority for proper handling and disposal procedures. All plastic, glassware, and solutions can be disinfected with chloramine T.

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

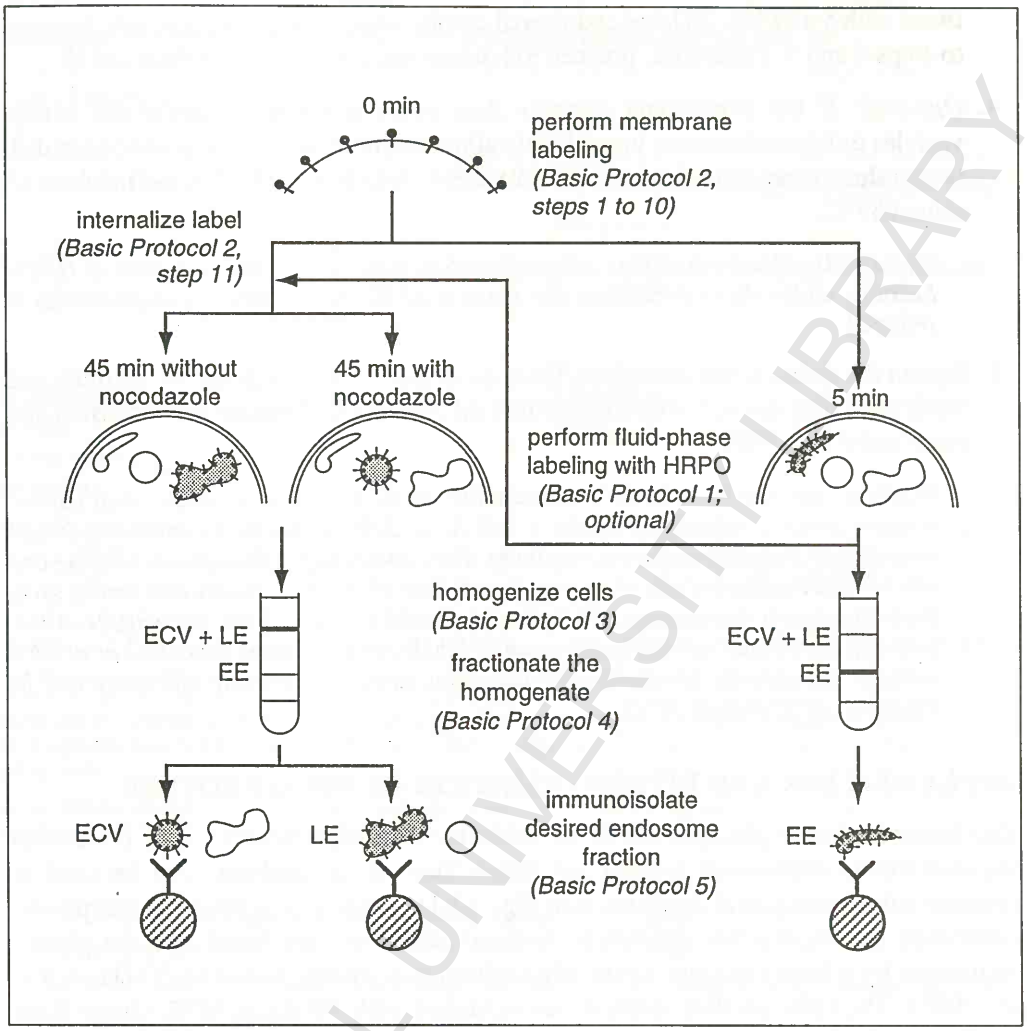


Figure 4.3.1 Outline of the complete procedure followed for the immunoisolation of early endosomes (EE), endosomal carrier vesicles (ECV), and late endosomes (LE). The G protein of the VSV virus is implanted into the plasma membrane by low pH–mediated fusion of the virus with the plasma membrane. Then, G protein molecules are cointernalized with horseradish peroxidase (HRPO) to introduce the G protein into EE, ECV, or LE, respectively. The cells are then homogenized and the PNS is prepared. The PNS is fractionated on a flotation gradient. The EE fraction or the ECV/LE fractions are used as input for immunoisolation using beads coupled to an antibody against the G protein cytoplasmic domain. Depending on the internalization conditions for the G protein, ECV can be separated from LE.

Wash the cells

1. Chill monolayer cultures of BHK-21 cells on the cold plate. Place the cold plate on a rocking platform.

   The dishes must lie perfectly flat on the plate. A rocking platform should be used for all incubations, particularly those with small volumes (e.g., <5 ml).

2. Rapidly aspirate the medium and rinse twice with 5 ml ice-cold PBS to ensure proper chilling of cells and removal of cellular debris.

Bind WGA to cells

3. Aspirate the PBS and add 2.5 ml ice-cold WGA. Incubate 1 hr on the cold plate.
4. Aspirate the medium and wash twice with 5 ml PBS.

ECU IASI CENTRAL UNIVERSITY LIBRARY

Bind VSV to cells

5. Aspirate the PBS and add 2.5 ml ice-cold VSV. Incubate 1 hr on the cold plate.

   Addition of 50 μg VSV protein to a 10-cm dish of confluent BHK-21 cells results in the fusion of 8.5 μg VSV protein with the plasma membrane; this corresponds to ~80 molecules of VSV G protein/μm² membrane surface area (Gruenberg and Howell, 1987).

6. Aspirate the virus solution and wash the cells twice with 5 ml PBS.

7. Remove the dish from the cold plate, aspirate the wash solution, and add 20 ml of 37°C fusion medium. Incubate 30 sec at 37°C.

   Low pH–mediated fusion at 37°C results in the incorporation of G protein molecules in the correct transmembrane orientation. Because all steps except the fusion are done on ice and because endocytosis is arrested at pH 4.9, G protein is restricted to the plasma membrane.

8. Aspirate the fusion medium and quickly place the dish back on the cold plate. Rinse twice with 5 ml ice-cold PBS.

Remove unfused VSV

9. Remove unfused virions by aspirating the PBS, then adding 5 ml of ice-cold 60 mM GlcNAc. Incubate 1 hr on the cold plate.

10. Aspirate the GlcNAc and wash twice with PBS.

Internalize G protein

11. Add 10 ml IM/BSA and incubate 5 min *or* 45 min at 37°C.

    After 5 min, G protein is found predominantly in early endosomes; after 45 min, it is found primarily in late endosomes. Endosomal carrier vesicles contain the G protein after 45 min if microtubules are depolymerized by treatment with nocodazole (see Basic Protocol 1).

12. *Optional:* If a general marker for endosomes is required, e.g., to establish a balance sheet for fractionation, cointernalize HRPO with the G protein (see Basic Protocol 1).

PREPARATION OF VSV STOCKS

SUPPORT PROTOCOL 1

For this immunoisolation experiment, the antigen is provided by the cytoplasmic domain of the transmembrane glycoprotein G of the enveloped virus, vesicular stomatitis virus (VSV). VSV stocks are produced and harvested as described in Gruenberg et al. (1989). Briefly, subconfluent 2-day cultures of BHK cells are infected with 0.2 pfu VSV/cell and incubated 19 to 20 hr at 37°C. Medium containing virus is removed from the cells and centrifuged 30 min at $7000 \times g_{avg}$ in an HS4 rotor to remove cell debris. The supernatant is layered on top of a two-step gradient of 10% and 55% (w/v) sucrose in 50 mM Tris·Cl/100 mM NaCl/0.5 mM EDTA (pH 7.4) and centrifuged 90 min at $90,000 \times g_{avg}$. Virions are collected at the 10%/55% interface. Aliquots of 50 and 200 μg are frozen in liquid nitrogen and stored at –70°C.

BCU IAS CENTRAL UNIVERSITY LIBRARY

BASIC PROTOCOL 3

HOMOGENIZATION OF TISSUE CULTURE CELLS

Gentle homogenization conditions should be used to limit possible damage to endosomal elements, particularly when using fluid-phase markers. Clearly, the markers should remain entrapped in vesicles (latent) after homogenization. In addition, harsh conditions should always be avoided to limit the breakage of lysosomes and consequent proteolysis by released hydrolases. Cells grown for 14 to 16 hr are easily homogenized, so each step of the homogenization process should be monitored using a phase-contrast microscope. A balance sheet should be used to follow the distribution of protein and markers during homogenization, flotation, and immunoisolation (see Table 4.3.2 for an example). Labeled cells are scraped from the dish, homogenized, and centrifuged to prepare a PNS.

Materials

Monolayer cultures of BHK-21 cells labeled with VSV G protein and/or fluid-phase marker (see Basic Protocol 1 or Basic Protocol 2)
PBS (*APPENDIX 2E*), 4°C
Homogenization buffer (HB; see recipe)
Protease inhibitor cocktail (see recipe)

15-ml plastic conical centrifuge tube
Refrigerated centrifuge and rotor appropriate for cell sedimentation
22-G needle attached to a 1-ml syringe
Phase-contrast microscope
Beckman Airfuge or equivalent ultracentrifuge

Additional reagents and equipment for assay of horseradish peroxidase activity (Support Protocol 2)

Collect cells

1. Remove cells from the dish by scraping with a rubber policeman in 2.5 ml of 4°C PBS and transfer to a 15-ml plastic conical centrifuge tube. Centrifuge 5 min at 700 × *g*, 4°C.

   Typically, the contents of three to five dishes are pooled into a 15-ml tube.

2. Remove and save the supernatant. Resuspend cell pellet very gently in 5 ml homogenization buffer using a wide-bore pipet. Centrifuge 10 min at 1500 × *g*, 4°C.

3. Analyze the supernatant for the presence of markers due to cell breakage during removal from the plate.

   Cell breakage during removal should not exceed 5% to 10%.

Homogenize cells

4. Add 0.5 ml protease inhibitor cocktail in HB to the cell pellet and resuspend using a 1000-µl pipettor with a blue (1000-µl) tip.

   Homogenization is easier at a relative high cell density, e,g, 20% to 30% (v/v).

   Cells grown in suspension or that are released from the dish by trypsin treatment may require harsher homogenization conditions. The use of a tight-fitting glass-glass Dounce homogenizer (15 to 20 passages of the pestle) or a "cell cracker" may be required to achieve good cell breakage. A cell cracker is a device used to pass cells through a precision clearance between the wall of a metal chamber and a metal ball bearing (Balch and Rothman, 1985).

5. Draw the homogenate three to ten times through a 22-G needle attached to a 1-ml syringe. Monitor homogenization using a phase-contrast microscope. Be careful not to over-homogenize.

PCU IASI CENTRAL UNIVERSITY LIBRARY

Nuclei should be intact and free of cellular materials.

Damage to endosomes and lysosomes can be quantified by measuring the latency of content markers, lysosomal enzymes, or other integral markers (latency corresponds to the percentage of a marker that remains entrapped within vesicles—e.g., endosomes) after homogenization. Under optimal conditions, latency should be >75% after homogenization.

6. Centrifuge the homogenate 15 min at 1500 × *g*, 4°C and collect the PNS and the nuclear pellet (NP). Resuspend the nuclear pellet in 1 ml PBS. Save a 50-µl aliquot of the PNS and the nuclear pellet for analysis.

   The PNS is ready for fractionation.

Determine latency of markers

7. Dilute a 20-µl aliquot of the PNS with HB to a total volume of 0.2 to 1 ml, depending on the amount of PNS necessary to detect the marker.
8. Centrifuge 15 min at 200,000 × *g* (e.g., 20 psi in a Beckman Airfuge), 4°C.
9. Collect the supernatant. Resuspend the pellet in the same volume of HB as used in step 7.
10. Determine the amount of HRPO or other marker in the original PNS, in the pellet, and in the supernatant (see Support Protocol 2).
11. Calculate the amount of marker in the pellet as a percentage of the total amount present in the PNS (yield).

DETERMINATION OF HORSERADISH PEROXIDASE ACTIVITY

SUPPORT PROTOCOL 2

Internalized HRPO is measured using *o*-dianisidine and peroxide as substrates; the brown-colored product is quantitated by measuring the absorbance at 455 nm with a spectrophotometer.

Materials

Subcellular fraction containing internalized HRPO (see Basic Protocol 3)
HRPO substrate solution (see recipe)
HRPO standards: 10 to 100 ng HRPO/ml in the same buffer used to prepare or dilute sample
10 µM potassium cyanide (KCN; optional)

1. Quickly mix 0.1 ml of sample with 0.9 ml substrate solution. Mix 0.1 ml of each HRPO standard with 0.9 ml substrate solution. Incubate at room temperature in the dark. Time the reaction.

   If necessary, sample may be diluted with PBS (APPENDIX 2E) or other buffer before analyzing an 0.1-ml aliquot.

2. When a brown color develops, note the time and measure the absorbance of sample and standards at 455 nm.

   With small amounts of HRPO, the reaction can be developed for up to 2 hr in the dark. It is not necessary to stop the reaction, but the heme group of HRPO can be blocked at the desired time by adding a respiratory inhibitor such as 10 µm KCN.

3. Use the values obtained for HRPO standards to quantitate the amount of HRPO in the sample.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

BASIC PROTOCOL 4

FLOTATION-GRADIENT FRACTIONATION OF TISSUE CULTURE HOMOGENATES

Early and late endosomal fractions can easily be prepared from postnuclear supernatants (PNS) by flotation in a step gradient of 40.6% to 8.6% (w/w) sucrose/D_2O. Early endosomes containing the transferrin receptor, the small GTPase rab5 protein, and annexin II can be separated from late endosomes containing the rab7 protein, lysosomal glycoproteins (lgps), and the cation-independent mannose-6-phosphate receptor (see Table 4.3.1). Fractions from this gradient can be used to determine the general protein composition of the fractions in high-resolution two-dimensional gels (*UNIT 10.4*) and to reconstitute the different endosomal membrane fusion events (see Fig. 4.3.2).

Materials

Postnuclear supernatant (PNS; see Basic Protocol 3)
Sucrose solutions (see recipes)
Homogenization buffer (HB): 250 mM sucrose/3 mM imidazole

Refractometer
SW60 centrifuge tubes
Ultracentrifuge and SW60 rotor
Peristaltic pump connected to a 50-µl capillary tube

1. Dilute the PNS 1:1 (v/v) with 62% sucrose solution in H_2O to give a solution of 40.6% sucrose/0.5 mM EDTA and check the final sucrose concentration with a refractometer.

2. Load 1 ml of the diluted PNS in the bottom of an SW 60 centrifuge tube and overlay sequentially with 1.5 ml of 16% sucrose solution in D_2O, then 1 ml of 10% sucrose solution in D_2O. Add HB to fill the tube.

   Early and late endosomes can also be isolated on a gradient with 35% and 25% sucrose solutions in H_2O. Also, the low-density sucrose step (10% in D_2O or 25% in H_2O) can be omitted to yield fractions containing mixed populations of early and late endosomes.

3. Mount the tubes in an SW60 rotor and centrifuge 1 hr at 14,000 × g_{avg} (35,000 rpm), 4°C.

4. In a 4°C cold room, use a peristaltic pump connected to a 50-µl capillary tube to collect the gradient.

   The early endosome fraction containing the transferrin receptor, the small GTPase rab5, and annexin II is collected from the 16%/10% interface (or 35%/25% interface). The late endosome fraction containing the rab7 protein, lysosomal glycoproteins

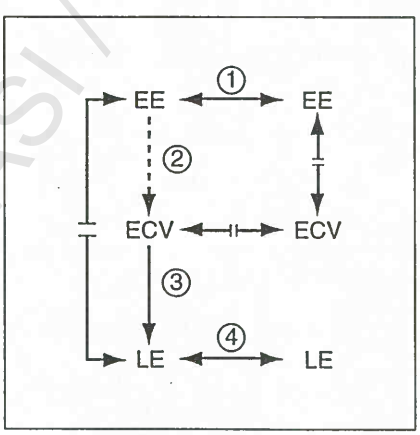


Figure 4.3.2 Outline of the different endosome fusion events reconstituted in an in vitro fusion assay. EE undergo homotypic fusion with each other (1; Gruenberg and Howell, 1989). Several components regulating this process have been identified (Gruenberg and Clague, 1992). ECV form from EE (2) in a process that requires vacuolar ATPase activity (Clague et al., 1994). ECV fuse with LE (3) in a process facilitated by microtubules and cytoplasmic dynein (Aniento et al., 1993a). LE also undergo homotypic fusion with each other (4). All other combinations (EE-ECV, ECV-ECV, and EE-LE) are refractory to fusion (Aniento et al., 1993a). Abbreviations: ECV, endosomal carrier vesicles; EE, early endosomes; LE, late endosomes.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

(lgps), and the cation-dependent mannose-6-phosphate receptor (CI-MPR) is collected from the upper portion of the 10% cushion (or 25% cushion in H_2O).

5. Divide early and late endosome fractions into 50-µl aliquots, then freeze and store them in liquid nitrogen.

   Aliquots should be placed in a liquid nitrogen bath for rapid freezing; similarly, they should be thawed as rapidly as possible prior to use or analysis.

IMMUNOISOLATION OF ENDOSOMAL FRACTIONS

BASIC PROTOCOL 5

Immunoisolation can be used to purify any subcellular fraction as long as a specific antibody is available. The specific antibody is coupled to magnetic or polyacrylamide beads with a covalently coupled species-specific linker antibody, then the beads are used to precipitate the desired organelle. A similar strategy has been used to immunoisolate other organelles with different antigen/antibody couples (Morré et al., 1988; Howell et al., 1989a). The procedure for immunoisolating endosomes using VSV G protein as the antigen is described here.

Materials

Rabbit anti-mouse immunoglobulin covalently coupled to polyacrylamide or magnetic beads (e.g., Immunobeads, Bio-Rad)
PBS/BSA: PBS (*APPENDIX 2E*) containing 2 mg/ml BSA, ice cold
Mouse antibody against the cytoplasmic domain of VSV G protein (P5D4; Kreis, 1986)
Endosomal fractions (see Basic Protocol 4)
Microcentrifuge, refrigerated
Low-speed rotating wheel (2 to 20 rpm)

1. Resuspend 1 to 10 mg rabbit anti-mouse immunoglobulin covalently coupled to polyacrylamide or magnetic beads in 1 ml PBS/BSA and microcentrifuge 2 to 3 min at 3000 × *g*, 4°C.

   Immunobeads are also available with antibodies to immunoglobulins from other species.

2. Wash the beads four times with PBS/BSA by centrifugation as in step 1.

3. Add 5 to 10 µg mouse anti–G protein antibody/mg bead in a total volume of 1 ml PBS/BSA for 1 to 10 mg of beads. Rotate end over end overnight at 4°C.

4. Remove excess antibody by washing three times with PBS/BSA. Resuspend the beads at 10 mg/ml in PBS/BSA.

5. Dilute the endosome fraction 1:3 with ice-cold PBS/BSA in a 1.5-ml microcentrifuge tube to a final volume ~1 ml.

6. Add 1 mg beads with bound antibodies to the cellular fraction. Rotate end over end 2 hr at 4°C.

   For either polyacrylamide or magnetic beads, 1 mg solid support (a pellet of ~20 µl) is used for immunoisolation from 300 to 500 µg PNS protein or 50 µg protein of a fraction collected from the gradient (the equivalent of cells from 10 to 20 cm^2 of tissue culture surface area).

7. Centrifuge the beads/fraction mixture 5 min at 3000 × *g* in a refrigerated centrifuge. Collect the supernatant for analysis. Gently resuspend the pellet in 1 ml PBS/BSA to prevent vesicularization and repeat the centrifugation.

B.C.U. „M. Kogălniceanu" IAŞI
CHIMIE

BCU IASI / CENTRAL UNIVERSITY LIBRARY

If the sample is to be analyzed by gel electrophoresis, the last wash should be in PBS without BSA.

The amount of immunoisolated protein cannot be quantitated directly because of the added BSA and antibodies. However, it is possible to obtain an estimate by using cells metabolically labeled with [^{35}S]methionine.

8. Resuspend the beads plus bound vesicles in 100 µl medium or buffer.

The choice of medium or buffer depends on how the sample will be analyzed; for example, use homogenization buffer (see recipe) for a cell-free assay of endosome fusion (Gorvel et al., 1991; Aniento et al., 1993a) or lysis buffer for two-dimensional gel electrophoresis (UNIT 10.4).

SUPPORT PROTOCOL 3

PREPARATION OF TISSUE CULTURE CELLS FOR SUBCELLULAR FRACTIONATION

Cells from a baby hamster kidney line (BHK-21) are grown as monolayers in G-MEM supplemented with 5% (v/v) fetal calf serum (FCS), 10% (v/v) tryptose phosphate broth, and 2 mM glutamine in a humidified 5% CO_2, 37°C incubator. The cells are seeded at 4 $\times 10^4$ cells/cm^2 in 10-cm tissue culture dishes (~3.2 $\times 10^6$ cells per dish) to form an even, confluent monolayer by 14 to 16 hr after seeding. These conditions yield ~1.3 $\times 10^7$ cells and ~2.5 to 3.0 mg total protein per dish at the time of the experiment. This protocol has the additional advantage that cells from 14- to 16-hr cultures can easily be homogenized. Typically, two to three 10-cm dishes are used per experimental condition, and fifteen to twenty 10-cm dishes are used for preparing stocks.

To isolate endosomal carrier vesicles, microtubules can be depolymerized prior to labeling and homogenization to inhibit transport of the marker from early endosomes to late endosomes and lysosomes. The conditions required to depolymerize microtubules may, however, vary between cell types. In BHK-21 cells, preincubating the cells 1 to 2 hr at 37°C in the presence of 10 µM nocodazole is sufficient (Gruenberg et al., 1989). No apparent depolymerization of BHK microtubules is observed after prolonged incubation at 4°C. In contrast, the depolymerization of MDCK microtubules requires both a cold treatment and an incubation at 37°C in the presence of nocodazole (Bomsel et al., 1990).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers see SUPPLIERS APPENDIX.

7.4 M medium

MEM containing:
10 mM TES [*N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid]
10 mM MOPS [3-(*N*-morpholino)propanesulfonic acid]
15 mM HEPES (*N*-2-hydroxyethylpiperzazine-*N*′-2-ethanesulfonic acid)
2 mM NaH_2PO_4
0.42 mM $NaHCO_3$
Adjust pH to 7.4
Store sterile medium 4 to 8 weeks at 4°C

Homogenization buffer (HB)

250 mM sucrose
3 mM imidazole, pH 7.4
10 mM Tris·Cl, pH 7.4 or 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid)
Store sterile buffer 4 to 8 weeks at 4°C

RSU IASI CENTRAL UNIVERSITY LIBRARY

HRPO substrate solution

0.05 M sodium phosphate buffer, pH 5.0
0.3% (v/v) Triton X-100
0.342 mM *o*-dianisidine
0.003% (v/v) H_2O_2
Store at 4°C protected from light

Prepare this solution in very clean glassware to avoid contamination. Add o-dianisidine first; after it has been solubilized, add H_2O_2. Discard if a straw yellow color develops.

Internalization medium

MEM containing:
5 mM D-glucose
10 mM HEPES (*N*-2-hydroxyethanepiperazine-*N'*-2-ethanesulfornic acid)
Adjust pH to 7.4
Store sterile medium 4 to 8 weeks at 4°C

Protease inhibitor cocktail

Stock solution
10 mM leupeptin
10 mg/ml aprotinin
1 mM pepstatin
Store several months at –20°C

Working solution: Dilute stock solution 1:1000 in homogenization buffer (see recipe).

Sucrose solutions

10% (w/w) sucrose solution in D_2O
0.303 M sucrose
3 mM imidazole, pH 7.4
Store sterile solution 4 to 8 weeks at 4°C. The refractive index of this solution should be 1.3478 at 20°C.

16% (w/w) sucrose solution in D_2O
0.497 M sucrose
3 mM imidazole, pH 7.4
Store sterile solution 4 to 8 weeks at 4°C. The refractive index of this solution should be 1.3573 at 20°C.

25% (w/w) sucrose solution in H_2O
0.806 M sucrose
3 mM imidazole, pH 7.4
Store sterile solution 4 to 8 weeks at 4°C. The refractive index of this solution should be 1.3723 at 20°C.

35% (w/w) sucrose solution in H_2O
1.177 M sucrose
3 mM imidazole, pH 7.4
Store sterile solution 4 to 8 weeks at 4°C. The refractive index of this solution should be 1.3904 at 20°C.

62% (w/w) sucrose solution in H_2O
2.351 M sucrose
3 mM imidazole, pH 7.4
1 mM EDTA
Store sterile solution 4 to 8 weeks at 4°C. The refractive index of this solution should be 1.4463 at 20°C.

BCU HART / CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

Subcellular fractionation was initially developed with rat liver and, in most cases, involves separation of organelles based on physical properties such as density. However, two major problems have interfered with the widespread use of subcellular fractionation in cell and molecular biology. First, studies of membrane transport and cellular organization have revealed that several subcellular compartments share similar physical properties and hence cofractionate, at least to some extent, in conventional gradients (e.g., Golgi and endosomes). In addition, the use of tissue culture cells is now widespread because cultures of cells can be manipulated in a manner impossible to achieve in animals or separate tissues. However, tissue culture cells are more difficult to fractionate than most tissues, presumably because of differences in organization of the cytoskeleton (Howell et al., 1989a).

Subcellular fractionation consists of two major steps: (1) disruption of the cellular organization (homogenization) and (2) fractionation of the homogenate to separate the different organelles. Since the early attempts to quantitatively fractionate liver tissue (Claude, 1946), methods have been developed to reproducibly disaggregate the tissue into a suspension of subcellular components and organelles called tissue homogenate (*UNIT 4.2*; DeDuve, 1955, 1971; Fleischer and Kervina, 1974; Beaufay and Amar-Cortesec, 1976). This homogenate can then be resolved by differential centrifugation into several fractions: (1) nuclei, heavy mitochondria, and plasma membrane; (2) light mitochondria, lysosomes, and peroxisomes; (3) Golgi, endosomes, and microsomes, and (4) cytosol. From these main fractions, organelles of interest can be further separated according to their size and density by additional purification steps. However, it is essential to point out that achievement of a complete purification is a barely accessible goal. Each type of organelle is composed of populations of elements that vary in size, density, and other properties on which separation relies. The complexity of tissues and cells, together with the structural alterations induced at all stages of fractionation, increases the physical heterogeneity. Additionally, the relative abundance of the organelle of interest in the cell will affect the yield in fractionation. Clearly, it is easier to purify major subcellular components (e.g., endoplasmic reticulum and mitochondria) than, for instance, transport vesicles, which are minor cellular elements. In the latter case, a low yield is often the price to be paid to achieve purification.

Subcellular fractionation

Traditional fractionation procedures are based on differences in the size and density of organelles or organelle-derived vesicles as measured in density gradients. Although sucrose is most widely used to form gradients, there are many other alternatives (e.g., Ficoll, Percoll, metrizamide, and Nycodenz). However, many of the smooth-membrane vesicles derived from the compartments involved in membrane traffic share common densities and are difficult to isolate by these procedures, so alternative methods have been developed. Table 4.3.2 lists a number of procedures that have been used to isolate endosomal fractions; see *UNIT 4.2* (and Morré et al., 1988 for a review). Gradients are extremely useful to separate populations of endosomes and lysosomes. However, gradients are often best combined with one of the techniques described below to achieve higher levels of purification. This unit describes the use of a flotation gradient combined with a method to immunologically isolate of endosomal fractions using the vesicular stomatitis virus (VSV) G protein.

Density-shift methods rely on altering the density of a specific compartment by the introduction of a perturbant. These methods have proven extremely useful for isolation of fractions from the endocytic/lysosomal pathway because of the ease with which density perturbants can be loaded into these compartments from the extracellular space. This experimental approach was pioneered by the DeDuve group, who used the detergent Triton WR-1339 to isolate lysosomes from rat liver (Leighton et al., 1968). Ligands coupled to horseradish peroxidase (Courtoy et al., 1984; Ajioka and Kaplan, 1986; Stoorgovel et al., 1987) or gold (Beardmore et al., 1987) have also been used to isolate endosomes. Golgi-derived clathrin-coated vesicles can be separated using a newly synthesized cholinesterase activity present in the vesicles (Fishman and Fine, 1987).

Free-flow electrophoresis separates subcellular compartments on the basis of charge and was first used to purify lysosomes from rat liver (Stahn et al., 1970). This method has been particularly useful for the separation of endosomes and lysosomes, which are deflected

PSU IAS/FERRA UNIVERSITY LIBRARY

Table 4.3.2 Methods of Subcellular Fractionation of Endosomes

Separation method	References
Density modification	
DAB shift procedure	Ajioka and Kaplan (1986); Courtoy et al. (1984); Stoorgovel et al. (1987)
Colloidal gold method	Beardmore et al. (1987); Beaumelle et al. (1990)
Free-flow electrophoresis	Marsh et al. (1987); Schmid et al. (1988)
Immunoisolation	Howell et al. (1989b); Thomas et al. (1992); Aniento et al. (1993a); Emans et al. (1993)
Gradients	
Percoll	Pertoft et al. (1978); Sahagian and Neufeld (1983); Bergeron et al. (1985)
Nycodenz	Kindberg et al. (1984); Mullock et al. (1989)
Sucrose	Khan et al. (1986); Mueller and Hubbard (1986); Gorvel et al. (1991); Aniento et al. (1993a); Emans et al. (1993)

towards the anode in the electric field. A two-step procedure in which endosomes and lysosomes are first separated from all other organelles by free-flow electrophoresis and then from each other by Percoll-gradient centrifugation has been carried out using both CHO and BHK-21 cells (Marsh et al., 1987). This work has been extended to separate two distinct populations corresponding to early and late endosomes (Schmid et al., 1988).

Immunoisolation relies on the unique localization of an antigen exposed on the outer surface of the compartment of interest, rather than on a physical property of the organelle (for a review see Howell et al., 1989b). Isolation is carried out using a specific antibody bound to a solid support (e.g., magnetic beads, cellulose fibers, acrylamide beads, or *Staphylococcus aureus* cells). The use of immunoisolation for separating different populations of endosomes is described in detail in Basic Protocol 5 and illustrated in Figure 4.3.1. This technique has been used to isolate endosomes using the cytoplasmic domain of the G protein as antigen (Howell et al., 1989a,b).

Different solid supports are available: heterodisperse polyacrylamide beads (Bio-Rad), *Staphylococcus aureus* expressing protein A (Pansorbin, Calbiochem), monodisperse magnetic beads (Dynal; Howell et al., 1989b), Eupergit particles (Röhm GmbH), and cellulose fibers (Richardson and Luzio, 1986; Luzio et al., 1988). The efficiency of immunoisolation for one particular antigen/antibody pair may vary with different immunoadsorbents depending on the abundance and/or accessibility of the antigen on the membrane of interest. Clearly, the actual immunoisolation experiment provides the only appropriate trial for selecting an immunoadsorbent. In the method described in this unit, the G protein of VSV is present at an estimated density of 80 G molecules/μm^2 membrane surface area (Gruenberg and Howell, 1987). The cytoplasmic domain of the G protein is relatively short (consisting of the 22 carboxyl-terminal amino acids) and the antibody used was raised against a synthetic carboxyl terminal peptide (Kreis, 1986). To reconstitute fusion in the cell-free assay, magnetic beads (Dynal) with a coupled linker antibody raised in sheep against the Fc domain of mouse IgG (Howell et al., 1989b) or anti-mouse polyacrylamide beads (Bio-Rad) were used. In this system >70% of fluid phase markers internalized for 5 to 45 min could be immunoisolated, by adding ≤300 μg of PNS protein to the beads; enrichment was 10- to 15-fold (Gruenberg and Howell, 1987) based on the ratio of markers immunoisolated on specific beads to markers bound nonspecifically to control beads.

The immunoisolated fractions immobilized on solid supports retain their fusion activity (Gruenberg and Howell, 1986, 1987), with

RUTGERS CENTRAL UNIVERSITY LIBRARY

only a minimal decrease in latency of fluid phase markers (Gruenberg et al., 1989). A significant advantage is that the bound fraction can be introduced into and then retrieved from a reaction mixture at the end of the reaction while bound to the solid support (Gruenberg and Howell, 1986, 1987, 1989; Emans et al.,1989). Retrieval is simple, and separation can be repeated sequentially with the same fraction through a series of experimental conditions. Separation is easily achieved with magnetic beads because they can be sedimented within minutes using a small magnet. However, a magnet should only be used to retrieve the beads after the immunoisolation step itself to limit magnetic aggregation of the beads.

By combining a flotation gradient with immunoisolation, it has been possible to purify early endosomes, endosomal carrier vesicles, and late endosomes (see Fig. 4.3.1). In the first case, the G protein was internalized for 5 min so it distributed in early endosomes. Then cells were homogenized and the PNS was fractionated on a flotation gradient. The early endosome band (interface between 16% and 10% sucrose in D_2O) was then used as input for immunoisolation. To isolate endosomal carrier vesicles (ECVs) and late endosomes, the G protein was internalized for 45 min, either in the absence of microtubules to introduce the G protein into ECVs or in the presence of microtubules to introduce the G protein into late endosomes (see Table 4.3.1 and Fig 4.3.1). After homogenization, the PNS was fractionated on the flotation gradient. In this case, the ECV/late endosome band (interface between 10% sucrose in D_2O and homogenization buffer) was used as input for immunoisolation. Thus, depending on the internalization conditions, either ECVs or late endosomes can be immunoisolated.

Analysis of subcellular fractions

The functional specialization of subcellular organelles is reflected by their morphology, protein composition, and enzymatic activities, all of which can be used to characterize and analyze purified cell fractions.

Markers

Traditionally enzymes that are localized, ideally in one cell organelle (e.g., sugar transferases in the Golgi), are used as marker enzymes for the corresponding compartments. Purification of an organelle can then be quantified by following the activity of the marker enzyme, if known, during the fractionation procedure (see example in Table 4.3.2).

The preparation of specific antibodies against proteins that localize to specific subcellular compartments has rendered biochemical and morphological characterization easier than in the past. In particular, the recently discovered large family of small GTPases, each of which has a restricted distribution, is becoming an essential tool for monitoring the composition of subcellular fractions (Zerial and Stenmark, 1993; Huber et al., 1994).

Exogenous markers can be endocytosed and thus serve as markers for endocytic organelles. The markers can be in the fluid phase or added to the plasma membrane.

Protein profiles

The protein patterns in high-resolution two-dimensional gels have been obtained for several subcellular compartments, allowing the identification of otherwise unknown proteins. In addition, the availability of two-dimensional databases (*UNIT 10.4*; Celis et al., 1994) is dramatically improving the criteria used for analysis of the protein composition of organelles.

Morphological analysis

Organelles retain many of their structural features after isolation. Therefore, electron microscopy is an important method for analyzing fractions. However, many compartments can have the same appearance—tubular or vesicular—so electron microscopy is best combined with immunocytochemistry or immunogold labeling of frozen sections.

Balance sheet

Typically, the distribution of the desired component, for example a new protein or the mutagenized form of a known protein, is compared with the distribution of known markers. Analysis of the purified fraction and assessment of the fractionation procedures requires a complete balance sheet of the values measured for the different markers that are being used, including a general marker (e.g., a protein), specific markers, and contaminants, if any. These values must be expressed so as to indicate the yield and enrichment of specific markers, as well as contaminants. The following parameters must be calculated for the different steps of the fractionation.

The **specific activity** measures the amount of the desired component in the fraction normalized to the total amount of protein. Traditionally, the specific activity refers to the activ-

ECU IASTI OLIVER - UNIVERSITY LIBRARY

Table 4.3.3 Balance Sheet for Isolation of Rat Liver Lysosomes[a]

	β-*N*-acetyl-D-glucosaminidase				
	Protein (g)	Total act. (units)	Spec. act. (U/mg)	Yield (%)	RSA
Homogenate	5.000	65.0	0.013	100.0	1.0
NM fraction	2.175	34.4	0.016	52.9	1.2
PMS fraction	2.825	30.6	0.011	47.1	0.8
L fraction	0.110	10.7	0.097	16.5	7.5
Fractions 1-3	0.011	6.9	0.623	10.6	143.9
Fractions 4-5	0.084	3.5	0.042	5.4	8.1

[a]Abbreviations: β-NAGase, β-N-acetylglucosaminidase; total act., β-NAGase total activity; Spec. act., β-NAGase specific activity; RSA, β-NAGase relative specific activity.

ity of a marker enzyme/amount of total protein, and is measured in units/mg protein. Clearly, the amounts of the desired component can also be quantified by other techniques, for example by immunoprecipitation with specific antibodies followed by SDS-PAGE, autoradiography, and scanning of the autoradiogram.

The **yield**, the activity of the marker enzyme in the fraction/total activity of the marker enzyme or protein in the starting material (i.e., homogenate), expressed as a percentage, is a measure of the amount of the material recovered in that fraction.

The **relative specific activity (RSA)** is the absolute specific activity of the fraction/absolute specific activity of the starting material. This value provides an adequate estimate for the degree of purification achieved in each fraction.

An example illustrating the use of these parameters after fractionation of rat liver lysosomes is presented in Table 4.3.3. Fractionation is followed using a spectrophotometric assay to quantify the enzymatic activity of a lysosomal enzyme, β-*N*-acetyl-D-glucosaminidase, as described (Aniento et al., 1993b). Livers excised from rats are homogenized, then the homogenate is centrifuged 10 min at 3000 × *g*, 4°C, to prepare a nuclear/mitochondrial pellet (NM fraction) and a postmitochondrial supernatant (PMS fraction). Fraction L corresponds to the pellet obtained after two sequential centrifugations of the PMS for 10 min at 25,000 × *g*, 4°C. The pellet is then resuspended in homogenization buffer and fractionated in a metrizamide gradient (Wattiaux et al., 1978; Aniento et al., 1993b). At the end of the centrifugation, fractions are collected and analyzed. Fractions 1 to 3 are pooled together and contain mainly lysosomes. Fractions 4 to 5 contain mainly mitochondria but also some lysosomes.

Functional characterization

The use of in vitro assays has become an essential tool to study complex cellular processes in the test tube, where they can be easily manipulated and analyzed. Most steps of membrane traffic have now been reconstituted in vitro, leading to the identification of several components regulating membrane transport (Rothman and Orci, 1992; Rothman and Warren, 1994). Using an in vitro assay to monitor endosome fusion, both early and late endosomes have been observed to exhibit a high homotypic fusion activity (Tuomikoski et al., 1989; Bomsel et al., 1990; Gorvel et al., 1991; Thomas et al., 1992; Aniento et al., 1993a; Emans et al., 1993). These observations suggest that both compartments are very dynamic in vivo, which is in agreement with video-microscopy observations in vivo (Hopkins et al., 1990). However, early endosomes do not undergo direct fusion with late endosomes in vitro (see Fig. 4.3.2). Intermediate endosomal carrier vesicles mediate transport from early to late endosomes in vivo and undergo fusion with late endosomes in vitro. These carrier vesicles do not exhibit homotypic fusion properties and do not undergo fusion with early endosomes, as expected. The interactions of carrier vesicles with late endosomes is specifically facilitated by the presence of polymerized microtubules, a process that requires the presence of the minus end–directed motor cytoplasmic dynein when using nonpolarized cell extracts (Aniento et al., 1993a). In contrast, in polarized MDCK cells, both cytoplasmic dynein and the plus

POLLARD CENTRAL UNIVERSITY LIBRARY

end–directed motor kinesin are required (Bomsel et al., 1990).

The fusion of endosomal carrier vesicles with late endosomes was also characterized by immunogold labeling of cryosections in the electron microscope. In these experiments, the endosomal carrier vesicles and late endosomes from two separate cell populations were specifically labeled with 5- and 16-nm gold particles, respectively, and the corresponding fractions were then prepared using a flotation gradient. These fractions were mixed in the fusion assay in the presence of cytosol and incubated at 37°C; finally, the mixture was processed for electron microscopy. When the assay was carried out in the absence of ATP to inhibit fusion, no mixing of 5- and 16-nm gold particles was observed, and only the 16-nm gold particles colocalized with antibodies against two lysosomal glycoproteins (4a1 and 2a5), as expected. In the presence of ATP, however, both types of gold particles were mixed within the same vesicular profiles, which also contained the 4a1 and 2a5 antigens. These experiments show that, upon fusion, the 5-nm gold particles originally present in endosomal carrier vesicles were transferred to late endosomes, which contained internalized 16-nm gold particles as well as the 4a1 and 2a5 antigens (Aniento et al., 1993a). It is therefore, proposed that early and late endosomes are separate but dynamic compartments connected by endosomal carrier vesicles. These vesicles are formed from early endosomes, then they are transported on microtubules to late endosomes, where they fuse (see Fig. 4.3.2; Aniento et al., 1993a). The process of ECV formation, in particular the inclusion of internal membranes into these so-called multivesicular structures, is still unclear. Recent data, however, show that an active vacuolar ATPase is required for ECV formation (Clague et al., 1994).

Critical Parameters

The major impediment to successful fractionation of tissue culture cells is the difficulty encountered when trying to produce an ideal homogenate, that is, the release of various subcellular components (i.e., organelles) as a free suspension of intact, individual elements. Often, it is necessary to compromise between complete homogenization and damage to cellular compartments. After homogenization, cytoplasmic aggregates that contain cytoskeletal elements as well as various organelles can often be observed. To some extent, these aggregates can form after homogenization when conditions are too extreme due to the presence, for instance, of DNA released by breakage of the nucleus. However, aggregates can also reflect some cellular organization, particularly due to the cytoskeleton, that remains after homogenization. Some organelles also remain associated with cytoskeletal elements surrounding the nucleus and become entrapped in "clumps" of cytoplasm that readily sediment. As a result, as much as 50% of the components of the homogenate may be pelleted along with the nucleus during the initial centrifugation step. Because the cytoplasmic organization of different tissue culture cell lines varies enormously, homogenization conditions must frequently be optimized for each cell line.

Homogenization of tissue culture cells

Under ideal homogenization conditions, particulate organelles such as nuclei, mitochondria, lysosomes, and peroxisomes remain intact. Golgi complexes, plasma membranes, and reticular organelles will fragment and can, at least to some extent, vesiculate. However, these fragments will retain many of the in vivo activities of the corresponding organelles (e.g., protein translocation across the endoplasmic reticulum membrane). Following homogenization, nuclei are totally and exclusively removed by centrifuging 10 min at 500 to 1000 × *g* to form a nuclear pellet (NP). The PNS contains cytosol and other organelles in free suspension. It is not possible to outline a general protocol suitable for the production of a "reasonable" homogenate for all tissue culture cells, because different conditions are required for different cell types and experimental purposes. However, four parameters influence the efficiency of homogenization of cultured cells and may provide general guidelines.

Growth and experimental conditions of the cells

To achieve comparable homogenization over extended periods of time, it is essential to use cells that are grown from the same original parent culture (frozen in most cases immediately after cloning). The growth conditions of the cells must also be standarized and optimized for homogenization to remain reproducible. Cells that have been growing in a confluent state for several days are often more difficult to homogenize, possibly because they have a more complex network of cytoplasmic filaments (Kreis and Vale, 1993). In many cases, better homogenization is achieved with

ECU IAS / CENTRAL UNIVERSITY LIBRARY

cells that are 12 to 36 hr post passage, depending upon the density at which cells are seeded. Finally, the experimental conditions to which cells are subjected (e.g., prolonged cold treatment) may also influence homogenization conditions. Cells growing in suspension (or cells normally grown in monolayer but brought into suspension, e.g., with trypsin) can require significantly harsher homogenization conditions than monolayers of cells that are mechanically scraped from the dish.

Homogenization medium

Upon homogenization, the cell components are unavoidably transferred from a physiological environment to an entirely different medium. The homogenization medium should fulfill a number of basic requirements: protect the organelles against osmotic burst and possible damage due to proteases, preserve their enzymic activities and biochemical functions, prevent any kind of agglutination, minimize leakage of constituents from organelles, and interfere as little as possible with biochemical analysis and functional assays. It is obvious that it is not always possible to fulfill all of these requirements, and the most appropiate medium will largely depend on the purpose of the experiment.

To maintain the structural integrity of particulate organelles, isotonic (250 to 320 mosM) conditions at neutral pH are preferable. These conditions are also essential to retain the content within the lumen of vesicles derived from the organelles of the secretory and endocytic pathways. Some buffers appear to be more useful for the production of a monodisperse homogenate in isotonic sucrose solutions. Among these are imidazole (Beaufay and Amar-Cortesec, 1976) and triethanolamine/acetic acid (the latter first introduced for and routinely used in free-flow electrophoresis experiments; Stahn et al., 1970). Some cells are very difficult to break in isotonic buffers. The forces required also disrupt the nuclear membrane releasing DNA, which causes the cytoplasm to clump and prevents further separation. In these cases, hypotonic conditions (e.g., 3 mM imidazole or 10 mM Tris·Cl) may become necessary to achieve reasonable homogenization, but isotonic conditions should be restored as soon as possible. It may be important to include a cocktail of proteolytic inhibitors in the homogenization medium; the choice of these is largely empirical (Gruenberg and Howell, 1986; Howell et al., 1989a; Gruenberg and Gorvel, 1992).

The use of disruptive agents to depolymerize filaments of the cytoskeletal network—microtubules, microfilaments, and intermediate filaments (Kreis and Vale, 1993)—may improve the quality of the homogenate obtained from tissue culture cells. Many interactions between the cytoskeletal network and organelles are functionally significant in vivo, but must be disrupted to allow separation of individual vesicles in the homogenate. However, depolymerization of the microtubules (e.g., with nocodazole) and actin filaments (e.g., with cytochalasin) does not result in significantly better homogenization in many cases, presumably because of the presence of intermediate filaments, which are extremely difficult to depolymerize.

The addition of salts may or may not be recommended. High salt conditions (≤1 M KCl) have been used to reduce vesicular aggregation and to prevent redistribution of luminal proteins released during homogenization of guinea pig pancreas (Scheele et al., 1978). In spleen homogenates, agglutination is avoided by the addition of 0.2 M KCl (Bowers et al., 1967). However, salts can promote aggregation in rat liver and FAO cells (Devaney and Howell, unpub. observ.). High salt has also been reported to promote the condensation of intermediate filaments, a condition which is likely to increase aggregation of organelles (Steinert et al., 1982). Finally, and most importantly, high salt concentrations cause the release of peripheral membrane proteins, which can be an advantage or a disadvantage, depending on the goal of the experiment.

Homogenization devices

When the conditions of the cells and the homogenization medium are optimized, equally efficient homogenization can be obtained by many methods: by glass/glass (Dounce) or glass/teflon (Potter-Elvehjem) homogenization vessels and pestles; by repeated pipetting with a pipet or a plastic tip of an automatic pipettor (Harms et al., 1988); or by passage through a small-gauge needle with a syringe (Gruenberg et al., 1989). According to our experience, cells grown in suspension are better homogenized in a "cell cracker," a device used to homogenize cells by passage through a precision clearance (5 to 20 μm ± 1 μm) between the wall of a metal chamber and a metal ball bearing (Balch and Rothman, 1985). The choice of homogenization device often depends on the volume of the original sample.

PROTEIN CENTRAL LIBRARY

Evaluation of the homogenate

The results of homogenization should be assessed by morphological and biochemical means. Using a phase-contrast microscope, it is possible to assess the extent of cellular disruption, the appearance of intact and clean nuclei (a simple method to evaluate a good homogenization), and the absence of cytoplasmic clumps, which often appear upon overhomogenization when the nuclei start to break and the DNA is released into the cytoplasm. The quality of the homogenate is mainly revealed by its biochemical properties. Traditionally, a marker for each of the major organelles is assayed to determine its distribution in the nuclear pellet (NP) and PNS (see Table 4.3.1). In addition, the structural integrity of the organelles may be monitored by measuring the structure-linked latency of marker proteins or enzymes. With endocytic compartments, the latency of a marker endocytosed in the fluid phase can be easily quantitated.

Anticipated Results

After subcellular fractionation, the purified cellular components are characterized according to several criteria, including the composition of the fraction with respect to known markers, general protein pattern, and morphological appearance. Early endosomes, endosomal carrier vesicles, and late endosomes from BHK cells have different characteristics as summarized in Table 4.3.1.

Early endosomes can be labeled with a fluid-phase marker (i.e., horseradish peroxidase, HRPO), which is internalized from the medium during a 5-min incubation at 37°C. In addition, early endosomes have both the small GTPase rab5 (Gorvel et al., 1991) and the calcium-lipid binding protein annexin II (Emans et al., 1993) associated with the cytoplasmic surface of their membranes. Endosomal carrier vesicles can be labeled, after microtubule depolymerization, using HRPO internalized for 5 min and then chased for 45 min at 37°C (Aniento et al., 1993a). Late endosomes are labeled with HRPO internalized and chased under the same conditions in the presence of polymerized microtubules. Late endosomes contain the cation-independent mannose-6-phosphate receptor (CI-MPR; Geuze et al., 1983; Griffiths et al., 1988) and the small GTPase rab7 (Chavrier et al., 1990), and are enriched in several lysosomal glycoproteins (lgps; Geuze et al., 1983; Griffiths et al., 1988), particularly the 4a1 and 2a5 proteins (Aniento et al., 1993a).

The three endosomal compartments have also been characterized with respect to their general protein composition by analyzing the immunoisolated fractions in high-resolution two-dimensional gels (Aniento et al., 1993a; Emans et al., 1993). Early endosomes contain 50 major proteins, many of which cannot be detected at later stages of the pathway (e.g., proteins recycling with the plasma membrane). Late endosomes also exhibit a complex composition with ~50 major proteins; these include unique proteins (proteins that are not abundant, if present at all, at earlier stages of the pathway), such as 4a1, 2a5, and other highly glycosylated acidic proteins. The distribution of the 4a1 and 2a5 proteins (120 and 45 kDa, respectively) has been confirmed by cryoelectron microscopy using monoclonal antibodies. Both exhibit the physical characteristics of lysosomal glycoproteins in other cells (Kornfeld and Mellman, 1989). In contrast, endosomal carrier vesicles do not appear to contain unique proteins; >90% of the proteins they contain are also detected in late endosomes, consistent with the view that ECVs are transient elements of the pathway and function as carrier vesicles.

At the morphological level, the three components have also different appearance in BHK cells. Early endosomes consist of tubulo-vesicular elements and cisternal regions, whereas carrier vesicles are typically spherical, with a diameter ~0.4 μm and a multivesicular appearance. Late endosomes also include tubular, cisternal, and multivesicular elements but are often larger, with a diameter of ~1 μm, and (in BHK cells) typically contain large translucent inclusions (Gruenberg et al., 1989; Aniento et al., 1993a). In addition, the fate of an electron-dense tracer (gold particles) can be followed by electron microscopy after internalization in vivo. The tracer appears first in early endosomes, then in endosomal carrier vesicles, and finally in late endosomes, which can be labeled with gold particles conjugated to 4a1 or 2a5 antibodies after immunogold labeling of frozen sections. When the microtubules are depolymerized, internalized gold particles do not reach the compartments that contain the 4a1 and 2a5 antigens (late endosomes) and remain in structures that show the typical morphology of endosomal carrier vesicles.

Time Considerations

Cells for fractionation have to be seeded 14 to 16 hr before the experiment. Fluid-phase internalization takes 60 to 90 min depending on the internalization conditions (see Table

BCU IAS CENTRAL LIBRARY

4.3.1), including washes before and after the internalization step. When microtubules have to be depolymerized, an additional 60 to 90 min are required. For endosomes that will be immunoisolated using the VSV G protein, ~3 to 4 hr are required for viral fusion before internalization and an overnight incubation to bind antibodies to the beads. Once the markers have been internalized (if required), cells have to be scraped from the dishes and homogenized. It takes ~1 hr to obtain the PNS. The flotation gradient is centrifuged for 1 hr. Fractions collected from the flotation gradient can be analyzed for protein and endocytosed marker and frozen in liquid nitrogen. It is convenient to aliquot the fractions in small volumes (i.e., 50 to 100 µl) that must be frozen quickly to limit damage to the organelles. These aliquots must also be thawed very quickly and shortly before use.

Literature Cited

Ajioka, R. S. and Kaplan, J. 1986. Intracellular pools of transferrin receptors result from constitutive internalization of unoccupied receptors. *Proc. Natl. Acad. Sci. U.S.A.* 83:6445-6449.

Aniento, F., Emans, N., Griffiths, G., and Gruenberg, J. 1993a. Cytoplasmic dynein-dependent vesicular transport from early to late endosomes. *J. Cell Biol.* 123:1373-1387.

Aniento, F., Roche, E., Cuervo, A., and Knecht, E. 1993b. Uptake and degradation of glyceraldehyde-3-phosphate dehydrogenase by rat liver lysosomes. *J. Biol. Chem.* 268:10463-10470.

Balch, W.E. and Rothman, J.E. 1985. Characterization of protein transport between successive compartments of the Golgi apparatus: Asymmetric properties of donor and acceptor activities in a cell-free system. *Arch. Biochem. Biophys.* 240:413-425.

Beardmore, J., Howell, K.E., Miller, K., and Hopkins, C.R. 1987. Isolation of an endocytic compartment from A-431 cells using a density modification procedure employing a receptor-specific monoclonal antibody complexed with colloidal gold. *J. Cell Sci.* 87:495-506.

Beaufay, H. and Amar-Cortesec, A. 1976. Cell fractionation techniques. *In* Methods in Membrane Biology (E.D. Korn, ed.) pp. 1-99. Plenum, New York.

Beaumelle, B.D., Gibson, A., and Hopkins, C. 1990. Isolation and preliminary characterization of the major membrane boundaries of the endocytic pathway in lymphocytes. *J. Cell Biol.* 111:1811-1823.

Bergeron, J.J.M., Cruz, J., Kahn, M.N., and Posner, B.I. 1985. Uptake of insulin and other ligands into receptor-rich endocytic components of target cells: The endosomal apparatus. *Annu. Rev. Physiol.* 47:382-403.

Bomsel, M., Parton, R., Kuznetsov, S.A., Schroer, T.A., and Gruenberg, J. 1990. Microtubule- and motor-dependent fusion in vitro between apical and basolateral endocytic vesicles from MDCK cells. *Cell* 62:719-731.

Bowers, W.E., Finkenstaedt, J.T., and DeDuve, C. 1967. Lysosomes in lymphoid tissue. I. The measurement of hydrolytic activities in whole homogenates. *J. Cell Biol.* 32:325-338.

Celis, J.E., Rasmussen, H.H., Olsen, E., Madsen, P., Leffers, H., Honore, B., Dejgaard, K., Gromov, P., Vorum, H., Vassilev, A., Baskin, Y., Xingdang, L., Celis, A., Basse, B., Lauridsen, J.B., Ratz, G.P., Andersen, A.H., Walbum, E., Kjaergaard, I., Andersen, I., Puype, M., vanDamme, J., and Vandekerckhove, J. 1994. The human keratinocyte two-dimensional protein database (update 1994): Towards an integrated approach to the study of cell proliferation and skin diseases. *Electrophoresis* 15:1349-1458.

Chavrier, P., Parton, R.G., Hauri, H.P., Simons, K., and Zerial, M. 1990. Localization of low-molecular-weight GTP-binding proteins to exocytic and endocytic compartments. *Cell* 62:317-329.

Clague, M., Urbe, S., Aniento, F., and Gruenberg, J. 1994. Vacuolar ATPase activity is required for endosomal carrier vesicle formation. *J. Biol. Chem.* 269:21-24.

Claude, A. 1946. Fractionation of mammalian liver cells by differential centrifugation. I. Problems, methods and preparation of extract. *J. Exp. Med.* 84:51-60.

Courtoy, P.J., Quintart, J., and Baudhuin, P. 1984. Shift of equilibrium density induced by 3,3′-diaminobenzidine cytochemistry: A new procedure for the analysis and purification of peroxidase-containing organelles. *J. Cell Biol.* 98:870-876.

Davoust, J., Gruenberg, J., and Howell, K.E. 1987. Two threshold values of low pH block endocytosis at different stages. *EMBO J.* 6:3601-3609.

DeDuve, C. 1971. Tissue fractionation. Past and present. *J. Cell Biol.* 50:20D-55D.

DeDuve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., and Appelmans, F. 1955. Tissue fractionation studies. 6. Intracellular distribution pattern of enzymes in rat liver tissue. *Biochem. J.* 60:604-617.

Emans, N., Gorvel, J.-P., Walter, C., Gerke, V., Griffiths, G., and Gruenberg, J. 1993. Annexin II is a major component of fusogenic endosomal vesicles. *J. Cell Biol.* 120:1357-1370.

Fishman, J.B. and Fine, R.E. 1987. A trans Golgi–derived exocytic coated vesicle can contain both newly synthesized cholinesterase and internalized transferrin. *Cell* 48:157-164.

Fleischer, S. and Kervina, M. 1974. Subcellular fractionation of rat liver. *Methods Enzymol.* 31:6-40.

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

Geuze, H.J., Stoorgovel, W., Strous, G.J., Slot, J.W., Bleekemolen, J.E., and Mellman, I. 1983. Sorting of mannose-6-phosphate receptors and lysosomal membrane proteins in endocytic vesicles. *J. Cell Biol.* 107:2491-2501.

Gorvel, J.-P., Chavrier, P., Zerial, M., and Gruenberg, J. 1991. Rab 5 controls early endosome fusion in vitro. *Cell* 64:915-925.

Griffiths, G., Hoflack, B., Simons, K., Mellman, I., and Kornfeld, S. 1988. The mannose-6-phosphate receptor and the biogenesis of lysosomes. *Cell* 52:329-341.

Gruenberg, J. and Clague, M. 1992. Regulation of intracellular membrane transport. *Curr. Opin. Cell Biol.* 4:593-599.

Gruenberg, J. and Gorvel, J.-P. 1992. In vitro reconstitution of endocytic vesicle fusion. *In* Protein Targetting: A Practical Approach (A.I. Magee and T. Wileman, eds.) pp. 187-216. Oxford University Press, Oxford.

Gruenberg, J.E. and Howell, K.E. 1986. Reconstitution of vesicle fusions occurring in endocytosis with a cell-free system. *EMBO J.* 5:3091-3101.

Gruenberg, J. and Howell, K. E. 1987. An internalized transmembrane protein resides in a fusion-competent endosome for less than 5 min. *Proc. Natl. Acad. Sci. U.S.A.* 84:5758-5762

Gruenberg, J. and Howell, K.E. 1989. Membrane traffic in endocytosis: Insights from cell-free assays. *Annu. Rev. Cell Biol.* 5:453-481.

Gruenberg, J., Griffiths, G., and Howell, K.E. 1989. Characterisation of the early endosome and putative endocytic carrier vesicles in vivo and with an assay of vesicle fusion in vitro. *J. Cell Biol.* 108:1301-1316.

Harms, E., Kern, H., and Schneider, J.A. 1980. Human lysosomes can be purified from diploid skin fibroblasts by free-flow electrophoresis. *Proc. Natl. Acad. Sci. U.S.A.* 77:6139-6143.

Hopkins, C.R., Gibson, A., Shipman, M., and Miller, K. 1990. Movement of internalized ligand-receptor complexes along a continuous endosomal reticulum. *Nature* 346:335-339.

Howell, K.E., Devaney, E., and Gruenberg, J. 1989a. Subcellular fractionation of tissue culture cells. *Trends Biochem. Sci.* 14:44-47.

Howell, K.E., Schmid, R., Ugelstad, J., and Gruenberg, J. 1989b. Immunoisolation using magnetic solid supports: Subcellular fractionation for cell-free functional studies. *Methods Cell Biol.* 31A:264-292.

Huber, L.A., Ullrich, O., Takai, Y., Lutcke, A., Dupree, P., Olkkonen, V., Virta, H., de Hoop, M.J., Alexandrov, K., Peter, M., Zerial, M., and Simons, K. 1994. Mapping of Ras-related GTP-binding proteins by GTP overlay and two-dimensional gel electrophoresis. *Proc. Natl. Acad. Sci. U.S.A.* 91:7874-7878.

Khan, M.N., Savoie, S., Bergeron, J.J.M., and Posner, B.I. 1986. Characterization of rat liver endosomal fractions: In vitro activation of insulin-stimulable kinase in these structures. *J. Biol. Chem.* 261:8462-8472.

Kindberg, G.M., Ford, T., Blomhoff, R., Rickwood, D., and Berg, T. 1984. Separation of endocytic vesicles in Nycodenz gradients. *Anal. Biochem.* 142:455-462.

Kornfeld, S. and Mellman, I. 1989. The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* 5:483-526.

Kreis, T.E. 1986. Micro-injected antibodies against the cytoplasmic domain of vesicular stomatitis virus glycoprotein block its transport to the cell surface. *EMBO J.* 5:931-941.

Kreis, T.E. and Vale, R.D. 1993. Guidebook to the Cytoskeletal and Motor Proteins. Oxford University Press, Oxford.

Leighton, F., Poole, B., Beaufay, H., Baudhuin, P., Coffey, J.W., Fowler, S., and DeDuve, C. 1968. The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with Triton WR-1339. *J. Cell Biol.* 37:482-513.

Luzio, J.P., Mullock, B.M., Branch, W.J., and Richardson, P.J. 1988. Immunoaffinity techniques for the purification and functional assessment of subcellular organelles. *Prog. Clin. Biol. Res.* 270:91-100.

Marsh, M., Schmidt, S.L., Kern, H., Harms, E., Male, P., Mellman, I., and Helenius, A. 1987. Rapid, analytical, and preparative isolation of functional endosomes by free-flow electrophoresis. *J. Cell Biol.* 104:875-886.

Morré, D.J., Howell, K.E., Cook, G.M.W., and Evans, W.H. 1988. Cell-free analysis of membrane traffic. *Prog. Clin. Biol. Res.* 270.

Mueller, S.C. and Hubbard, A.L. 1986. Receptor-mediated endocytosis of asialoglycoproteins by rat hepatocytes: Receptor-positive and receptor-negative endosomes. *J. Cell Biol.* 102:932-942.

Mullock, B.M., Branch, W.J., vanSchaik, M., Gilbert, L.K., and Luzio, J.P. 1989. Reconstitution of an endosome-lysosome interaction in a cell-free system. *J. Cell Biol.* 108:2093-2099.

Pertoft, H., Warmegard, B., and Hook, M. 1978. Heterogeneity of lysosomes originating from rat liver parenchymal cells. Metabolic relationship of subpopulations separated by density-gradient centrifugation. *Biochem. J.* 174:309-317.

Richardson, P.J. and Luzio, J.P. 1986. Immunoaffinity purification of subcellular particles and organelles. *Appl. Biochem. Biotechnol.* 13:133-145.

Rothman, J.E. and Orci, L. 1992. Molecular dissection of the secretory pathway. *Nature* 355:409-415.

Rothman, J.E. and Warren, G. 1994. Implications of the SNARE hypothesis for intracellular membrane topology and dynamics. *Curr. Biol.* 4:220-233.

Sahagian, G.G. and Neufeld, E.F. 1983. Biosynthesis and turnover of the mannose 6-phosphate receptor in cultured Chinese hamster ovary cells. *J. Biol. Chem.* 258:7121-7128.

FCU IAS CENTRAL UNIVERSITY LIBRARY

Scheele, G.A., Palade, G.E., and Tartakoff, A.M. 1978. Cell fractionation studies on the guinea pig pancreas. Redistribution of exocrine proteins during tissue homogenization. *J. Cell Biol.* 78:110-130.

Schmid, S.L., Fuchs, R., Male, P., and Melman, I. 1988. Two distinct populations of endosomes involved in membrane recycling and transport to lysosomes. *Cell* 52:73-83.

Stahn, R., Maier, K.P., and Hannig, K. 1970. A new method for the preparation of rat liver lysosomes. Separation of cell organelles of rat liver by carrier-free continuous electrophoresis. *J. Cell Biol.* 46:576-591.

Steinert, P.M., Zachroff, R.V., Anyardi-Whitman, M., and Goldman, R.D. 1982. Isolation and characterization of intermediate filaments. *Methods Cell Biol.*. 24:399-419.

Stoorgovel, W., Geuze, H.J., and Strous, G.J. 1987. Sorting of endocytosed transferrin and asialoglycoprotein occurs immediately after internalization in HepG2 cells. *J. Cell Biol.* 104:1261-1268.

Thomas, L., Clarke, P., Pagano, M., and Gruenberg, J. 1992. Inhibition of membrane fusion in vitro via cyclin B but not cyclin A. *J. Biol. Chem.* 267:6183-6187.

Tuomikoski, T., Felix, M., Doree, M., and Gruenberg, J. 1989. Inhibition of endocytic vesicle fusion in vitro by the cell-cycle control protein kinase cdc2. *Nature* 342:942-945.

Wattiaux, R., Wattiaux-De Coninck, S., Ronveaux-Dupal, M.F., and Dubois, F. 1978. Isolation of rat liver lysosomes by isopycnic centrifugation in a metrizamide gradient. *J. Cell Biol.* 78:349-368.

Zerial, M. and Stenmark, H. 1993. Rab GTPases in vesicular transport. *Curr. Opin. Cell Biol.* 5:613-620.

Contributed by Fernando Aniento and Jean Gruenberg
University of Geneva Sciences II
Geneva, Switzerland

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Desalting, Concentration, and Buffer Exchange by Dialysis and Ultrafiltration

Various steps in protein purification result in concentrations of low-molecular-weight materials, such as $(NH_4)_2SO_4$, that exceed levels appropriate for subsequent procedures. In cases where simple dilution cannot sufficiently reduce the concentration of these impurities, dialysis or diafiltration—dialysis by ultrafiltration at constant volume—should be considered. Ultrafiltration is also convenient for concentrating protein samples without concomitantly concentrating low-molecular-weight impurities. This is achieved by forcing water and dissolved low-molecular-weight substances through a semipermeable membrane under pressure; the high-molecular-weight protein components are retained by the membrane. Alternatively, ultrafiltration can be carried out via a rapidly recirculating flow of solution over a semipermeable membrane (tangential-flow ultrafiltration). Dialysis and ultrafiltration are also effective methods for buffer exchange—i.e., the replacement of a buffer used for a previous procedure with a different buffer appropriate for a subsequent one.

Gel filtration (*UNIT 8.3*) can also be used for desalting, but because this purification technique normally employs a column containing a differentially porous gel matrix, rather than a semipermeable membrane, it is less readily adapted to large sample volumes. However, despite the availability of membranes with a wide range of porosities, dialysis and ultrafiltration are generally less efficient than gel filtration for separation of proteins on the basis of size difference, except where the difference in size is quite large (i.e., by a factor of ≥10; see Background Information).

This unit includes a variety of dialysis and ultrafiltration techniques that can be used for desalting, concentration, or buffer exchange. Basic Protocol 1 describes standard dialysis by diffusion across cellulose tubing, a technique used for desalting or buffer exchange. Basic Protocol 2 describes ultrafiltration under pressure, which can be used either for concentrating protein or, where the sample volume is replenished with a desired buffer, for desalting/buffer exchange (i.e., diafiltration). Alternate Protocol 1 describes ultrafiltration by tangential flow, which serves the same purposes as ultrafiltration under pressure but is better suited to handling large volumes of solution. Alternate Protocol 2 is an ultrafiltration technique employing centrifugal microconcentrators, which allow concentration of very small volumes of solution.

Additional information on standard dialysis using membrane diffusion is provided in *APPENDIX 3B*.

DESALTING AND BUFFER EXCHANGE BY DIALYSIS THROUGH REGENERATED CELLULOSE TUBING

BASIC PROTOCOL 1

When sample volume is <200 ml, dialysis can be accomplished by placing the sample in standard dialysis tubing immersed in a buffer of the desired ionic strength and continuously stirring the buffer. For larger volumes, multiple dialysis bags may be used, but desalting or buffer exchange on a large scale is more rapidly achieved by tangential-flow diafiltration (see Alternate Protocol 1). Because the exposure time of the sample to the buffer is relatively long, the composition, pH, and ionic strength of the buffer should be dictated by considerations of protein stability; when large volumes are involved, buffer price may be a factor as well.

Copyright © 1995 by John Wiley & Sons, Inc.

BOURNEMOUTH UNIVERSITY LIBRARY

Materials

Dialysis tubing: regenerated cellulose, MWCO 12,000 (e.g., Spectra/Por Type 4, Spectrum) for general protein use or MWCO 3500 (e.g., Spectra/Por Type 3) for small proteins, stored at 4°C in plastic bag
50% (v/v) ethanol in large wash bottle
10 mM EDTA, pH 8.0
0.05 M $NaHCO_3$
Protein solution (clarified by centrifugation if necessary)
Appropriate dialysis buffer

1-liter beaker
Tubing closures, two (one weighted and one not weighted) per dialysis bag (e.g., Spectra/Por closures and weighted closures, Spectrum)
Small plastic funnel
Conductivity meter and probe (optional)

Prepare the dialysis tubing

1. Select tubing of appropriate MWCO that has a width such that the sample can be contained in sections ≤25 cm in length (see Table 4.4.1). Cut the dry tubing to the proper length, allowing an additional 10 cm for closing off the ends.

   Handle tubing only with powder-free gloves to avoid getting skin oils on membrane.

2. Immerse the sections of tubing in distilled water and rub each between two fingers until the layers separate. Using a large wash bottle containing 50% ethanol, rinse the inside and outside of each section thoroughly.

   The ethanol rinse serves to remove glycerol (used as a humectant) and residual sulfides (from the manufacturing process).

3. Mix 400 ml of 10 mM EDTA (pH 8.0) and 400 ml of 0.05 $NaHCO_3$ in a 1-liter beaker. Transfer the wetted and washed tubing to this beaker and stir 30 min using a magnetic stirrer.

4. Decant the EDTA/$NaHCO_3$ solution and replace with 800 ml distilled water. Stir 10 min on the magnetic stirrer, decant the water, and repeat the water wash. Transfer the tubing to fresh distilled water.

   If not to be used promptly, the tubing should be stored covered at 4°C in water containing 0.1% sodium azide. Once wetted, the membranes must not be allowed to dry out.

   Boiling tubing is not recommended prior to use in protein work. For more vigorous removal of glycerol and sulfide in situations where these impurities could be troublesome, the tubing may be immersed in 1 liter of 50% ethanol and heated 1 hr at 50°C,

Table 4.4.1 Approximate Capacities for Dialysis Tubing of Various Diameters

Flat width (mm)	Inflated diameter (mm)	Volume/length (ml/cm)
10	6.4	0.3
18	11.5	1.0
25	16	2.0
45	29	6.4
54	34	9.3
75	48	18

BAU IASI / CENTRAL UNIVERSITY LIBRARY

instead of being rinsed as in step 2. Regenerated cellulose tubing may be autoclaved, but this alters the MWCO. Individually packaged 14-cm lengths of radiation-sterilized tubing containing no glycerol, sulfides, or heavy metals are available at several MWCO values between 8,000 and 25,000 (Spectra/Por Biotech Sterile Membranes, Spectrum). These are suitable for small (2- to 20-ml) samples.

Dialyze the sample

5. Close one end of a washed section of tubing firmly with a weighted tubing closure. Squeeze the tubing between two fingers to remove excess water from inside.

   Alternatively, the tubing section may be closed by tying two simple overhand knots near one end.

6. Slip the open end of the tubing section over the narrow end of a small plastic funnel, then add the protein sample through the funnel. When sample addition is complete, remove the funnel and work the sample down the tubing by squeezing with the fingers to eliminate air space. Clamp the end shut with a non-weighted tubing closure (or knot the tubing carefully twice; see annotation to step 5) and trim off excess tubing. Check for leakage at both ends by applying moderate pressure to the sample area.

   Excessive air space will lead to dilution of the sample by buffer that enters and displaces air during dialysis; however, extra space should be allowed if the sample is very high in salt concentration, as it may be necessary to permit an increase in volume to prevent rupture of the bag. If the sample is very viscous, it should be diluted with the buffer against which it will be dialyzed, to facilitate outward diffusion.

7. Place the filled dialysis bag in a beaker containing a volume at least ten times the sample volume of dialysis buffer at the desired temperature, making sure that the bag is completely immersed. Stir the dialysis buffer continuously with a magnetic stirrer.

   Continual stirring serves to maximize the difference in salt concentration between the interior and immediate exterior of the dialysis bag. A larger initial buffer-to-sample ratio (≥100:1) will speed attainment of the desired final ionic strength with fewer changes of dialysis buffer. Although diffusion rate is somewhat reduced in the cold, almost all protein dialysis is conducted at 4°C to enhance protein stability.

8. Follow the progress of the dialysis by measuring the conductivity of the dialysis buffer (if this is practical). Replace the used dialysis buffer with an equal volume of fresh buffer when the increase in conductivity slows down (indicating approach to equilibrium and a drop in the efficiency of the dialysis). Continue replacing the dialysis buffer in this way until the conductivity of the buffer remains essentially unchanged after stirring for 1 to 2 hr.

   If a conductivity meter is not available, the buffer should be changed at least twice, allowing 4 to 6 hr between changes to permit the system to approach equilibrium.

   Conductivity measurements are particularly useful when large amounts of salt are to be removed and when the buffer-to-sample ratio is <100:1. In such cases, several replacements of dialysis buffer will be necessary to achieve the desired final ionic strength; conductivity measurements permit dialysis to be completed in minimum time by eliminating guesswork regarding when the system is approaching equilibrium. When a conductivity meter is not used and large samples (or tubing diameters >20 mm) are involved, it may be necessary to extend the time between buffer changes to 8 to 12 hr to insure that equilibrium is attained.

9. Remove the dialysis bag from the buffer and rinse the exterior with distilled water. While holding the bag just below the upper closure (or knot), remove that closure (or cut off the knot). Carefully tip the bag and pour the dialyzed sample into an appropriate container. Confirm that the desired salt concentration was achieved by

BOULDER CENTRAL UNIVERSITY LIBRARY

comparing conductivity measurements of the dialyzed sample and unused dialysis buffer.

If the bag is extremely taut as a result of buffer infiltration, do not cut the upper knot; instead put one end of the tubing into a beaker large enough to hold the entire sample, then puncture the tubing near the lower knot with a long needle and allow the sample to drain directly into the beaker.

BASIC PROTOCOL 2

CONCENTRATION OR DIALYSIS BY ULTRAFILTRATION THROUGH ASYMMETRIC MEMBRANE DISKS

Protein samples at an initial concentration ≤20 mg/ml can easily be reduced in volume using an ultrafiltration system and membrane, as in the protocol described below. The stirred-cell system used in this protocol—in which nitrogen gas pressure forces molecules smaller than the MWCO through the pores—is available with a variety of chamber sizes and membrane properties. Because most salts pass through the membrane along with water, the sample volume is reduced without a significant change in its surrounding buffer composition, an obvious advantage over lyophilization.

Ultrafiltration can also be used to remove or exchange salts without changing the protein concentration. This process, called diafiltration, resembles dialysis but is driven by the rate of ultrafiltration (see Critical Parameters and Troubleshooting), not by a concentration gradient. Diafiltration is carried out as in the concentration protocol here, except that continuous or periodic additions of buffer of the desired final pH, concentration, and composition are made to compensate for changes in protein concentration that would otherwise occur as a result of volume reduction.

Materials

Protein solution (clarified by centrifugation if necessary)
N_2 gas cylinder and two-stage pressure regulator capable of controlling delivery up to 100 psi
5% ethanol

Ultrafiltration membrane (YM10, Amicon; PLGC, Millipore; or Molecular/Por Type C, Spectrum) with MWCO >10,000 and diameter to fit stirred cell
Ultrafiltration unit, stirred-cell type (Series 8000, Amicon, or Molecular/Por, Spectrum) with:
- Tubing for connection to N_2 source
- Glass funnel with stem diameter to fit through cell fill hole (optional; used in certain cell models)
- Ultrafiltrate collection container (Griffin beaker with graduations)
- Pressurized reservoir (optional; for use when sample size is large and protein is dilute, or when diafiltration is performed with constant readdition of new buffer): e.g., Spectra/Chrom Pressure Reservoir (Spectrum) or RG5 Fiberglass Reservoir (Amicon)

1. Rinse a new ultrafiltration membrane with distilled water to wet it thoroughly.

Handle membrane only with powder-free gloves and hold by the edge to avoid scratching the smooth, thin "skin" surface.

Some ultrafiltration membranes can be used more than once. The Millipore PLGC (MWCO 10,000) or PLTK (MWCO 30,000) cellulose membranes are intended mainly for single use. Many other suppliers indicate that their membranes can be reused until clogged or damaged. Such membranes may be more economical, but only if each user makes sure that the previously used membrane is undamaged and properly cleaned. Amicon YM series or Spectrum Molecular/Por Type C membranes can be cleaned by soaking in 0.01 M SDS solution or 1% Terg-A-Zyme (e.g., Baxter or VWR) for several

PUU IASI / CENTRAL UNIVERSITY LIBRARY

hours; the membranes are then thoroughly rinsed with distilled water before reuse. Used, cleaned membranes should be stored covered with 5% ethanol in a closed container at 4°C. For membranes other than the ones mentioned, the manufacturer's instructions must be consulted for a suitable cleaning procedure.

2. Load the membrane into the stirred cell with the "skin" or glossy side up. Complete assembly of the ultrafiltration system, taking special care to ensure that O ring seals (if used) are not nicked and are properly placed on the membrane disk to prevent leakage around the membrane when pressure is applied.

   The size and thus the membrane diameter of the ultrafiltration cell should be as large as practical for the amount of sample to be concentrated; a larger cell reduces the necessity to refill the unit and provides the largest filtration area, speeding the concentration process. Cells are commonly available in sizes ranging from 10 to 400 ml.

3. Fill stirred cell with sample to the maximum operating volume, either by pouring the sample directly through the fill hole or by using the glass funnel provided with the cell in systems where the fill hole is small. Replace the fill-hole cover and check that the pressure-release valve (often an integral part of the fill-hole cover) is in the closed position.

   If a funnel is used, contact between the funnel tip and membrane surface should be avoided.

4. Place the unit on a magnetic stirrer and stir at the maximum rate that does not generate foam or a sizeable vortex at the surface of the solution. Direct the outlet tubing to the ultrafiltrate collection container.

   Because of protein stability requirements, most ultrafiltrations are conducted in the cold. It is advisable to check the surface temperature of the magnetic stirrer periodically; if it is warm, a styrofoam sheet 0.5-in. thick should be placed between the stirrer surface and the cell bottom.

5. Connect the cell to the nitrogen gas regulator and pressurize to 50 psi.

   Flow of liquid into the outlet tubing should quickly become evident. The system should be examined for appearance of liquid at the base of the cell where the cylinder body contacts the O ring; this indicates a leak around the membrane, in which case the unit should be depressurized and its assembly examined.

   Lower or higher pressures may be needed depending on the membrane type and MWCO value, but pressure should not exceed 70 psi in most stirred cells.

6. Monitor the flow rate and ultrafiltrate volume until the desired rate of concentration has been achieved.

   If the total sample volume exceeds the cell capacity, the cell should be refilled when the volume has decreased by one-third. Sample additions are then continued until all of the sample has been added to the cell. A pressurized sample reservoir is recommended when the sample volume is many times the cell capacity.

   At a constant temperature and pressure, the flow rate is a function of the filter area and the degree to which concentration polarization can be avoided (see Critical Parameters). Vigorous stirring is needed to minimize concentration polarization of protein at the membrane surface. Buildup of protein on the surface will result in slow filtration, even when the protein concentration of the sample is relatively low. Filtration rates at 4°C are often only one-half those seen at 25°C because of the influence of viscosity.

   Do not concentrate the sample to the point where it is no longer in contact with the stir bar, as lack of stirring will lead to severe concentration polarization in all but the most dilute samples. Overconcentration makes sample recovery difficult and may require readdition of buffer to wash the membrane, thereby adding to the volume.

BCM INSTITUTE CENTRAL UNIVERSITY LIBRARY

7. Turn off the stirrer and the nitrogen flow to the cell. Slowly release the internal cell pressure, then turn the stirrer back on and allow stirring to continue for 2 min to resuspend protein in the gel layer. Finally, turn the stirrer off and remove the concentrated sample by pouring or pipetting through the fill hole.

   If a pipet is used and the membrane is of the reusable type, care should be taken not to touch the pipet to the membrane surface. An alternative means of removing the sample is to unclamp the cell and pour out the sample, after checking that the body of the unit is firmly screwed into the base.

8. If the membrane is of the reusable type, hold it by the edges and rinse it well with distilled water. Proceed as soon as possible with the manufacturer's suggested cleaning procedure, keeping the membrane wet at all times, then store the cleaned filter in 5% ethanol in a closed container at 4°C.

ALTERNATE PROTOCOL 1

DIAFILTRATION OR CONCENTRATION BY TANGENTIAL-FLOW ULTRAFILTRATION

The most rapid method for dialyzing large samples is the use of ultrafiltration membranes fabricated into hollow fibers or spiral-wound layers, or, as in this protocol, in the form of stacked plates. The sample is allowed to flow tangentially over these membranes with recirculation. The semipermeable nature of the membrane then permits water and materials with molecular weights that are substantially below the MWCO value to cross through the membrane surface relatively unimpeded; retained materials, including the proteins of interest, are returned to the sample reservoir. To use this technique for diafiltration (as opposed to concentration), fresh dialysis buffer of the desired composition should be added to the sample reservoir in a volume equal to that of the ultrafiltrate collected, so that the concentration of macromolecular species remains unchanged. Removal of microsolutes to which the membrane is completely permeable can be estimated from the volume of ultrafiltrate (or diafiltrate, in this case) collected. Collection of an amount of diafiltrate equal to 3 vol of original sample will remove 95% of the starting solute to which the membrane is permeable, while a diafiltrate of 5 vol will correspond to a 99% separation.

The same process is used to concentrate samples, except that the retentate volume is allowed to decrease until the desired degree of concentration is reached. Maintaining a rapid flow over the membrane with a moderate transmembrane pressure minimizes concentration polarization and keeps the filtration rate high even as protein concentration in the sample increases.

The procedure described here is based on the Millipore Minitan acrylic ultrafiltration system, which employs stacked filter plates and a peristaltic pump. It is designed for sample sizes from 200 ml to 2 liters, but will work with larger samples if they are dilute. Equivalent units include the Spectrum MP-3 system and Amicon CH2PRS concentrator, both of which employ spiral-wound membrane cartridges.

***Additional Materials** (also see Basic Protocol 1)*

1% (w/v) Terg-A-Zyme detergent (e.g., Baxter or VWR), 40°C
0.05% (w/v) sodium azide

Minitan membrane filter plates (Millipore) made of cellulose (PLGC, MWCO 10,000 or PLTK, MWCO 30,000) or polysulfone (PTTK, MWCO 30,000)
Minitan acrylic ultrafiltration system with pump assembly and tubing (Millipore)
Hose clamps (adjustable with screws)
Container capable of holding entire sample

DDU IASRI CENTRAL UNIVERSITY LIBRARY

Prepare the system

1. Select a filter plate of appropriate MWCO.

   Filtration rates are higher with MWCO 30,000 membranes, but the large volumes to be filtered in diafiltration may lead to some loss of proteins with molecular weight <50,000 using such membranes. Membranes of lower MWCO are also available (MWCO 1000, 3000, and 5000) but these filter more slowly and can reject some of the buffer ions, thus requiring more filtrate volume to provide complete buffer exchange.

2. Disassemble the Minitan cell to expose the lower acrylic manifold. Place the first silicone retentate separator on the acrylic surface. Carefully slip a filter plate down on the alignment pins until it is seated against the silicone separator. Follow this with the second silicone retentate separator, then another filter plate, until the desired number of filter plates have been stacked—varying the left-to-right orientation of plates at each level to permit tangential flow across each plate. After the last plate and separator have been inserted, replace the upper acrylic manifold, then the top steel plate. Complete the assembly by adding the washers and nuts, tightening them against the plate with a torque wrench to the recommended pressure.

   Layers of filter plates alternating with silicone separators can be added up to a total of 10, each with an area of 60 cm^2. Use multiple filter plates if available, especially when initial volumes exceed 2 liters. Additional filter plates increase the filtration rate in a linear fashion by providing more cross-sectional area.

3. Attach tubing to the retentate inlet on the lower manifold and run it through the pump to the sample reservoir. Likewise, attach tubing to the retentate outlet on the upper manifold and bring it to the sample reservoir. Attach a hose clamp on this section of tubing to permit the transmembrane pressure to be increased if necessary. Attach tubing from both filtrate outlets to a collection vessel; provide each section of outlet tubing with a hose clamp.

   A Y-shaped tube connector can be used to direct the output of both filtrate exit ports to a single piece of tubing.

4. Place a small amount of dialysis buffer in the sample reservoir and slowly bring the pump to a mid-range operating speed. Run the pump until the dialysis buffer has circulated through and purged the filter plates and all tubing lines. Stop the pump and discard the contents of the sample reservoir and collection vessel.

Perform ultrafiltration

5. Fill the sample reservoir with the protein sample to be diafiltered or concentrated. Restart the pump and adjust the flow to 100% recirculation by clamping off the filtrate outlets. Measure the flow and readjust the pump speed until the maximum cross-flow that provides an inlet pressure of 25 psi is achieved.

6. Open the filtrate outlets and begin collection of the ultrafiltrate.

   If filtrate flow is inadequate, some increase can be achieved by partially restricting the retentate outlet line by adjusting the screw on the hose clamp, thereby increasing the transmembrane pressure. The inlet pressure should not be allowed to exceed the recommended level for the tubing being used (usually 25 psi for silicone tubing with 4 mm i.d.). However, reduced tangential flow can lead to greater concentration polarization and ultimately a reduced dialysis rate.

   A recirculation flow/filtrate flow ratio of 20:1 may be satisfactory for dilute samples, but higher ratios are more successful at preventing concentration polarization. Because shear forces are normally not high during recirculation, protein stability is seldom reduced as a consequence of increasing this ratio.

BOKU VIBT / CENTRAL LIBRARY

7. As filtration proceeds, frequently refill the sample reservoir with fresh dialysis buffer to restore the original volume. Continue doing this until the volume of ultrafiltrate is three to five times the volume of the original sample, indicating that removal of diffusible material is 95% to 99% complete. If concentration is desired, allow the sample volume to decrease after this point until the desired volume is achieved.

8. Clamp off the filtrate line and slow the pump to terminate filtration. Remove the retentate inlet line from the sample reservoir and place it in a solution of fresh dialysis buffer, then allow the pump to push sample into the reservoir until a volume of buffer equal to the internal volume of the system has been collected (usually ~25 ml). Finally, stop the pump and remove the contents of the sample reservoir.

Clean the system

9. Clean the unit and membrane before disassembly by first pumping water through all lines to flush out remaining sample and buffer salts, then circulating 1% Terg-A-Zyme, 40°C, through the system for 30 min. Collect output from filtrate and retentate line in the same reservoir so that all solution is reused. Flush the system with 5 to 10 liters of water, discarding all output from the filtrate and retentate lines and making sure that all lines are flushed.

10. Disassemble the unit and store the filter plate(s) in 0.05% sodium azide in a covered plastic container at 4°C.

ALTERNATE PROTOCOL 2

MICROSCALE CONCENTRATION AND DESALTING WITH CENTRIFUGAL ULTRAFILTERS

Sample preparation for analytical purposes and small-scale purification may require concentration and/or desalting of volumes <1 ml. Stirred-cell assemblies (see Basic Protocol 2), although able to deal with amounts only slightly greater than that, often have unacceptably large internal volumes making sample recovery difficult. Moreover, as concentration proceeds the risk of excessive volume reduction (even to dryness) increases, because filtration with stirred cells is not self-limiting.

The advent of commercially available, single-use ultrafiltration microconcentrators powered by centrifugal force has eliminated many of these problems. These are available with a range of MWCO values in sizes suitable for handling volumes of 50 µl to 2 ml (e.g., Microcon units from Amicon). The microconcentrators can often be combined with an additional microporous membrane of 0.22-µm pore size to remove large particulates—e.g., polyacrylamide—at the same time the protein sample is being concentrated or desalted (e.g., Micropure Separator inserts from Amicon, used with the Microcon units). Centrifugal ultrafilters are also produced in sizes that will accommodate samples of ≥10 to 25 ml, but their limited membrane area and minimal remixing can result in low rates of concentration as compared with stirred cells, except when samples are dilute. Some recent designs appear to have greatly reduced concentration polarization (e.g., Centriprep from Amicon); these mostly involve improved methods of producing movement of liquid through the membrane via centrifugal force.

Additional Materials (also see Basic Protocol 1)

Centrifuge microconcentrator units: e.g., Microcon-10, -30, or -50 (MWCO 10,000, 30,000, and 50,000 respectively; Amicon); Ultrafree-MC with MWCO 5,000, 10,000, or 30,000 (Millipore); or Centri/Por with MWCO 10,000, 25,000, or 50,000 (Spectrum)

1. Pipet ≤500 µl of protein solution into microconcentrator and seal with the cap supplied by the manufacturer. When using a single concentrator or when multiple ones are not

DAU IASRI CENTRAL UNIVERSITY LIBRARY

matched in weight, prepare a balance tube (if necessary) by pipetting a volume of water equal to that of the sample into a previously used microconcentrator.

2. Place the microconcentrator opposite its balance tube in a microcentrifuge in a 4°C refrigerator or cold room. Microcentrifuge 20 to 40 min at the speed specified for the filter unit by the manufacturer.

A nonrefrigerated microcentrifuge may develop temperatures detrimental to protein samples when operated for extended periods; therefore it is usually best to have the microcentrifuge in a refrigerator or cold room for this operation, even though the filtration rate is reduced by the cold.

Microcon units can withstand forces of 13,000 × g, whereas the Ultrafree units are limited to 5,000 × g. If the sample is not sufficiently concentrated after 20 to 40 min, centrifugation should be continued an additional 20 min to see if further volume reduction occurs. Most centrifugal microconcentrators are designed to stop at a volume of 5 to 15 µl. Typical centrifugation times to concentrate 500 µl of 0.1% bovine serum albumin to 10 µl at 4°C are 25 min for the Microcon-30 and 70 min for the Microcon-10. These times are twice those observed at 25°C; recovery is ~95% in either instance.

If desalting without concentration is required, remove the microconcentrator from the microcentrifuge at the end of the centrifugation period, unscrew the sample reservoir from the filtrate tube, then pour off the filtrate from the filtrate tube. Replace the reservoir on the filtrate tube and add a sufficient amount of suitable buffer to restore the original sample volume. Replace the cap and invert to mix. If the sample does not show the desired final ionic strength as determined by conductivity measurement, repeat the centrifugation and readdition of buffer.

3. Once the desired degree of concentration has been achieved, remove the cap from the unit and place the filter portion upside down into the retentate tube supplied by the manufacturer. Place the reversed unit back in the microcentrifuge and spin 15 sec to drive the concentrate down to the tip of the tube. Remove the unit from microcentrifuge and recover the sample from the retentate tube.

With some microconcentrators, adsorption of protein to the walls of the unit as well as to the filter itself can be significant when the sample is very dilute (≤25 µg/ml). Passivation of the surfaces by soaking the entire unit overnight in 5% Tween-20, followed by thorough washing with distilled water, may help alleviate this problem. Suelter and Deluca (1983) suggest that modification of the protein solvent by addition of glycerol or Triton X-100 (below the critical micelle concentration) is superior to treatment of the surfaces of the microconcentrator.

COMMENTARY

Background Information

Used primarily for removal of excess salts or organic solvents (e.g., acetone or ethanol) employed as precipitants in protein purification or analysis, molecular separations by means of differentially permeable membranes have been an integral part of protein biochemistry since the earliest purifications. Membranes of natural origin have long been supplanted for use in dialysis by regenerated cellulose tubing. Relatively recent developments in membrane technology have provided the asymmetric designs now found in ultrafiltration membranes. This asymmetry refers to the "sidedness" built into ultrafiltration membranes, wherein a dense layer or skin ~0.1 to 0.5 µm in thickness with a carefully controlled pore size is bonded to one side of a thicker structural layer (composed of the same or different material) that has a large open-celled structure highly permeable to most molecules. A variety of materials have been used to fabricate these semipermeable membranes, ranging from cellulose and cellulose esters to polyethersulfone or polyvinylidene difluoride (PVDF). Because of proprietary interests, specific details concerning construction of these filters—particularly in regard to the more complex materials used in tangential-flow ultrafiltration—are not widely disseminated by their manufacturers. However, several

B. U. C. CHILE

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE LIBRARY

companies (Millipore, Amicon, and Spectrum) make available extensive information on their product lines and offer considerable technical advice and literature pertaining to their use.

All membranes, whether used for dialysis or ultrafiltration, are characterized by their molecular-weight cutoff (MWCO) value. This is usually defined as the molecular weight of a solute that is 90% prevented from penetrating the membrane under a chosen set of conditions. For some membranes the MWCO value is based on calibration using a mixture of dextrans, while for others it is based on calibration with specific globular proteins. How readily a particular protein is rejected by the membrane is a function of the shape, hydration state, and charge of the protein molecule. Hence, it is possible for two proteins of similar molecular weight to have decidedly different retention properties when a membrane with a stated MWCO that is near their molecular weight is used to separate them. Moreover, MWCO values are not sharp; rather, there is a gradual increase in retention as the size of solute molecules approaches and exceeds the average membrane pore size. Only at the point where all pores are smaller than a particular solute molecule is that molecule completely excluded (Saltonstall, 1992a).

In standard dialysis procedures, the requirement is for a large amount of a diffusible material to migrate through the membrane while high-molecular-weight components are retained. The rate of diffusion is influenced primarily by four factors: the concentration gradient of the diffusible material, the diffusion constant of that material, the area of the membrane, and the temperature. An increase in any or all of these variables produces more rapid dialysis; however the two that can most readily be varied are the gradient and the membrane area. The gradient should be kept large by rapid stirring of the dialysis buffer, which keeps the concentration of diffusible solutes as low as possible at the membrane exterior surface, as well as by frequent changes of the buffer solution. It is also desirable to stir the contents of the dialysis bag—which speeds up internal diffusion and minimizes concentration polarization—but this is not easily done without special equipment. The primary reason that dialysis is commonly used is that it is technically uncomplicated; even though the technique is moderately slow, arrangements intended to increase efficiency (e.g., rocking dialyzers) are rarely used (McPhie, 1971).

The rate of movement for permeable solutes in dialysis is difficult to predict, but the end state that can be achieved can be readily estimated. Ignoring possible inequality between the concentrations of low-molecular-weight ions on the two sides of a membrane that may result from a Donnan effect, the equilibrium salt concentration is predictable from simple dilution considerations. This is illustrated by the following example. If a 20-ml protein sample in 2 M ammonium sulfate/50 mM potassium phosphate is dialyzed against 1000 ml of a dialysis buffer containing 50 mM potassium phosphate without ammonium sulfate, the dilution factor will be $^{1,020}/_{20}$; thus, at equilibrium the $(NH_4)_2SO_4$ concentration will be 0.039 M. If, after equilibration, this buffer is replaced with 1000 ml of fresh 50 mM ammonium sulfate, upon equilibration the ammonium sulfate concentration will be <0.001 M. If only 200 ml of 50 mM potassium phosphate is used for each of the two buffer changes, however, the ammonium sulfate concentration in the sample will approach 0.016 M. It is thus highly desirable to change the dialysis buffer several times or to use a large ratio of dialysis buffer to sample. Both of these strategies serve to maximize the concentration differential that drives the process of dialysis towards the solute concentration present in the original dialysis buffer.

Separation by means of dialysis or ultrafiltration of two protein species that differ in their molecular weights (and R values; see Critical Parameters) is possible under some conditions, but generally will be incomplete unless the size difference is large. Saltonstall (1992b) has concluded that an efficient separation from an initial 1:1 ratio to a final 10:1 ratio would require a ratio of retention values of 10:1 (e.g., 0.9 versus 0.09), corresponding to a molecular-weight ratio of 23:1 for several different types of membrane.

Critical Parameters and Troubleshooting

Membrane impurities

Dialysis membranes made from regenerated cellulose generally contain appreciable amounts of heavy metals, with iron, zinc, and lead present at levels in excess of 1 to 10 ppm. They also contain ~0.1% sulfur from the manufacturing process. Pretreatment may be necessary to remove these materials when their presence is detrimental. Cellulose ester mem-

branes do not contain these contaminants and thus can be used directly; however they are less tolerant of organic solvents and high pH than regenerated cellulose membranes.

Ultrafiltration

The advantage of desalting processes based on ultrafiltration over those based on simple dialysis is that the rate of low-molecular-weight solute removal is not determined by a concentration differential, but rather by the flow rate of solvent and the rejection of the solute by the ultrafiltration membrane employed. The rejection (R) of a solute is defined by the relationship $R = 1 - (C_F/C_I)$, where C_F is the final concentration of the solute (in the filtrate) and C_I is the initial concentration of the solute (in the solution prior to filtration). In desalting by diafiltration, where the filtrate is replaced by fresh dialysis buffer at the same rate at which it is removed, a membrane is usually selected whose MWCO is much larger than the molecular weight of the solute to be removed. This produces an R equal to zero for the permeable solute, and, optimally, an R essentially equal to 1 (i.e., total rejection) for the protein(s) of interest—a situation in which loss of protein is minimized. When the diafiltration involves removal of a solute with an R equal to zero, the degree of washout of the solute can be calculated from the volume of filtrate collected relative to the original sample volume according to the equation: $\ln (C_I/C_R) = (V_F/V_I)$, where C_R is the final concentration of the solute in the retentate, V_F is the volume of ultrafiltrate collected (also equal to the volume of dialysis buffer added back to the sample), and V_I is the initial sample volume (which remains constant in diafiltration because of readdition of buffer). Using this equation, the volume of filtrate that must be collected relative to sample size (V_F/V_I) can be estimated for any degree of desalting required. Hence, to reduce the concentration of solute in the sample 10-fold (i.e., $C_I/C_R = 10$), the volume of filtrate collected must be 2.3 × the original sample volume; for a 100-fold reduction, this volume ratio must be 4.6.

Membranes for ultrafiltration are generally selected on the basis of the MWCO needed to retain the protein of interest but allow the maximum amount of other materials to pass through. It is usually best to choose an MWCO value that is roughly one-half the molecular weight of the species to be retained. This provides a reasonable margin of retention whereby almost none of the protein of interest should be lost, but at the same time provides the largest difference between the MWCO value and the molecular weight of the salts to be removed, thereby maximizing filtration rate. Table 4.4.2 lists flow rates for several types of membrane disks.

Concentration polarization

The term concentration polarization describes the concentration gradient of solutes (mainly those to which the membrane is not readily permeable) that forms immediately above the membrane on the side under pressure. This condition exists to some degree even at low solute levels. Concentration polarization can be reduced by efficient remixing of the solution—either by stirring action (in stirred-cell systems) or tangential movement of solution over the membrane (in tangential-flow systems). However, if the concentration of material within the gel layer at the membrane surface reaches a value that is sufficiently high, the filtration rate becomes pressure-independent and may fall to near zero.

One of the chief advantages of tangential-flow ultrafiltration systems is that they allow better control of concentration polarization than stirred-cell systems. In tangential-flow systems, concentration polarization is controlled by varying the recirculation rate of the retentate and adjusting the transmembrane pressure. Flux of filtrate through the membrane is a linear function of transmembrane pressure, until gel formation on the surface becomes the limiting factor. When this happens, increasing the velocity across the membrane surface without further increasing the transmembrane pressure will improve flux.

Adsorption of protein to the membrane

In regard to the degree of nonspecific adsorption of protein to membranes, some important differences exist among the various materials used in membrane fabrication. Most suppliers indicate that their cellulose membranes exhibit the least nonspecific adsorption of protein; these are therefore recommended for protein work even though they may have slower filtration rates than membranes fabricated from other materials. The actual amount of protein bound depends on the quantity present, its charge characteristics, and the membrane area. Even with the best membranes, losses of 1% to 5% are not uncommon when dealing with total quantities of protein in the range of 1 to 10 mg using a filter with a 43-mm diameter. The nature of the buffer can also affect adsorption of protein; some membranes

UNIVERSITY LIBRARY

Table 4.4.2 Typical Flow Rates for Disk-Type Ultrafiltration Membranes

Membrane[a]	MWCO (kD)	Approximate flow rate at 25°C (ml/min/cm^2)	
		For water	For protein solution
Amicon			
YM3	3	0.07[b]	0.07[c]
YM10	10	0.2[b]	0.15[c]
YM30	30	0.9[b]	0.2[c]
PM10	10	2.5[b]	0.2[c]
PM30	30	4.0[b]	0.3[c]
Spectrum			
Molecular/Por Type C	3	0.1[d]	—
Molecular/Por Type C	10	0.2[d]	—
Millipore			
Ultrafree	10	0.05[e]	0.05[f]
Ultrafree	30	0.4[e]	0.1[f]

[a]Addresses and telephone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.
[b]Determined at 55 psi.
[c]Determined at 55 psi using 0.1% albumin.
[d]Determined at 50 psi.
[e]Determined at 2000 × *g*.
[f]Determined at 2000 × *g* using 0.1% IgG.

exhibit altered flow properties when high levels of ions are present. In this regard, phosphate buffers seem to present more of a problem than Tris buffers.

Precipitation of protein

Desalting sometimes results in precipitation of a portion of the protein in the sample. Although the insoluble material can be removed by centrifugation, it should not be discarded until assays establish that the desired protein has remained in solution. If the precipitated material includes the protein of interest, it can often be redissolved in a solution of higher ionic strength.

Anticipated Results

When carried out correctly, dialysis and diafiltration should produce a nearly complete recovery of the protein of interest in a buffer solution of the desired type and ionic strength. In the case of dialysis, this requires that the procedure be carried out for a sufficient time period to equilibrate the salt concentrations inside and outside the dialysis bag. In the case of diafiltration with an asymmetric membrane, this requires that a volume of filtrate be collected that is equal to approximately five times the initial sample volume.

The degree of concentration to be achieved by ultrafiltration should be that required for subsequent work. However, it should be kept in mind that once a protein concentration >30 mg/ml has been reached, further concentration is attained with increasing difficulty and production of very concentrated solutions can therefore be slow. Recovery of sample following concentration is generally >95%; failure to achieve this value usually indicates leakage into the filtrate or nonspecific binding to the membrane and/or concentration apparatus itself.

Time Considerations

Dialysis through regenerated cellulose tubing for desalting or buffer exchange usually requires 6 to 12 hr for volumes ≤10 ml and 12 to 24 hr for volumes from 10 to 200 ml. Dialysis time can be reduced by more frequent buffer changes, by increasing the ratio of dialysis buffer to sample, by increasing the surface area of the bag, and by monitoring the conductivity of the dialysis buffer to determine when the desired ionic strength has been reached.

The rate of ultrafiltration, whether performed for concentration or diafiltration, varies considerably as a function of the MWCO and surface area of the membrane and the initial

PU IAS / CENTRAL UNIVERSITY LIBRARY

sample concentration. A 10-fold reduction in sample volume (or 10-fold reduction in the concentration of a low-molecular-weight impurity by diafiltration) can usually be accomplished in 1 to 4 hr.

Literature Cited

McPhie, P. 1971. Dialysis. *Methods Enzymol.* 22:23-32.

Saltonstall, C.W. 1992a. Separations with dialysis and ultrafiltration: Theoretical and practical considerations. *Am. Biotechnol. Lab.* 10:32-34.

Saltonstall, C.W. 1992b. Dialysis and ultrafiltration: Estimating quantitative separations. *Am. Biotechnol. Lab.* 10:18-20.

Suelter, C.H. and Deluca, M. 1983. How to prevent losses of protein by adsorption to glass and plastic. *Anal. Biochem.* 135:112-119.

Key References

Pohl, T. 1990. Concentration of proteins and removal of solutes. *Methods Enzymol.* 182:68-83.

A recent review of the topics covered here.

Scopes, R.K. 1994. Protein Purification: Principles and Practice, 3rd ed. pp. 14-21. Springer-Verlag, New York.

Covers concentration and dialysis in the context of protein purification techniques.

Contributed by Allen T. Phillips
Pennsylvania State University
University Park, Pennsylvania

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Selective Precipitation of Proteins

Selective precipitation of proteins (Rothstein, 1994) can be used as a bulk method to recover the majority of proteins from a crude lysate, as a selective method to fractionate a subset of proteins from a protein solution, or as a very specific method to recover a single protein of interest from a purification step. Except for antibody-mediated precipitation (see *UNIT 10.8*), selective precipitation methods are usually not protein-specific; the process depends on the physical and/or chemical interaction of the precipitating agent with one or more proteins that possess certain characteristics. It may be necessary to use a combination of selective precipitation techniques to isolate the desired protein. The success of precipitation is monitored by measuring total protein content (*UNIT 3.4*), by identifying proteins in the fractions using SDS-PAGE (*UNIT 10.1*), and by performing a suitable bioassay to identify the fraction containing the desired protein.

Developing a selective precipitation technique for a particular protein requires identification of the appropriate precipitating agent and optimization of the precipitation procedure using that agent. Often, more than one reagent will successfully precipitate a particular protein, so selecting a reagent is a matter of identifying the one that provides the desired protein in the most optimal state. Some reagents cause denaturation or adversely affect bioactivity, and some form complexes that bind the protein tightly. Still others lack the ability to selectively enrich for the desired protein. This unit describes a number of methods suitable for selective precipitation (see Table 4.5.1). In each of the protocols that are outlined, the physical or chemical basis of the precipitation process, the parameters that can be varied for optimization, and the basic steps for developing an optimized precipitation are described.

Table 4.5.1 Protein Precipitative Techniques

Technique	Protocol
Salting out	Basic Protocol 1, Alternate Protocol 1
Isoionic precipitation	Basic Protocol 2, Alternate Protocol 2
Two-carbon (C_2) organic cosolvent precipitation of proteins	Basic Protocol 3
C_4 and C_5 organic cosolvent precipitation, phase partitioning, and extraction of proteins	Basic Protocol 4
Protein exclusion and crowding agents (neutral polymers) and osmolytes	Basic Protocol 5
Synthetic and semisynthetic polyelectrolyte precipitation	Basic Protocol 6
Metallic and polyphenolic heteropolyanion precipitation	Basic Protocol 7
Hydrophobic ion pairing (HIP) entanglement ligands	Basic Protocol 8
Matrix-stacking-ligand coprecipitation	Basic Protocol 9
Di- and trivalent metal cation precipitation	Basic Protocol 10

BOURNEMOUTH UNIVERSITY LIBRARY

Contributed by Rex E. Lovrien and Daumantas Matulis

Current Protocols in Protein Science (1997) 4.5.1-4.5.36

Copyright © 1997 by John Wiley & Sons, Inc.

Clarification of Samples

If raw samples are laden with particulate matter that cannot be centrifuged (at 5 to 50 × *g*) or filtered out by ordinary means, sometimes suspensions can be clarified by filtration on a 2- to 4-mm layer of "filter-aid" material—e.g., Filter-Aid (Schleicher and Schuell)—on Whatman no. 1 filter paper. Several kinds of filter aid materials for particulate and gelatinous samples are available—e.g., cellulosic fibers (Filter-Aid from Schleicher and Schuell or Hyflo Supercel from Fisher), diatomaceous minerals (e.g., Celatom from Aldrich, kieselguhr from Baxter, or diatomaceous earths from Sigma), silicates (e.g., Celite from Aldrich), specialized clay absorbents, calcium minerals, bentonite, or Fuller's earth. These should be experimented with—e.g., by adding 0.05 to 0.2 g of the material to 1 to 2 ml of the sample to find out which filter aid adsorbs gelatinous or particulate material without excessive adsorption of the protein. The layer or pad of filter aid for use with the main sample is prepared by adding a few grams to the filter paper and distributing it evenly by puddling it with water while pulling moderate suction on a Buchner funnel.

Frequently, a means for freeing proteins in solution from particles and haze by centrifugation or simple filtration can be quickly worked out by adjusting the pH or salt content (or both). The pH may be adjusted up or down by one to three pH units; also, salt may be added to make particulates and colloidal material easier to centrifuge or filter out. Bioactivity assays should be carried out even at the clarification stage to ensure that pH changes and salt additions did not destroy the activity or functions of the proteins to be isolated. Figure 4.5.1 summarizes the steps in clarification of protien solutions.

If the conservative techniques for clarification (filter aids and pH/salt variation) do not work, more aggressive techniques may be used that employ some of the rather powerful particle and colloid agglomerating agents such as synthetic polyelectrolytes (described below; see Basic Protocol 6), or metal-ion flocculating agents such as aluminum ions. These need to be added in quite limited amounts (in micrograms to a few milligrams per ml of sample), cautiously and slowly, to coagulate unwanted particles without precipitating sought-for proteins at this preliminary clarification stage. An advantage of metal ions as particulate coagulants or filter aids is that any excess of metal ions can be conveniently removed by conventional cation-exchange resins in K^+, Na^+, or NH_4^+ form (see Basic Protocol 2).

Unwanted particulates like whole cells, cell membranes, and haze may also be removed by centrifugation in some samples. However, for very large volumes (and where centrifugation is not successful) filtration using filter-aid compounds is frequently useful.

Pressure membrane ultrafiltration (*UNIT 4.4*) with a large MWCO (>30,000) may be tried if such membranes pass the sought-for proteins and hold back the haze particles.

Concentration of Samples

Very dilute samples beyond the reach of normal precipitating agents for proteins may need to be concentrated before adding precipitating agents. Lyophilization (freeze-drying) is sometimes successful for concentrating certain proteins that are able to withstand removal of virtually all water. However, a majority of proteins, including clotting-cascade and many blood plasma and cytosol proteins, cannot withstand lyophilization. It is more generally preferred, therefore, to use pressure membrane ultrafitration (*UNIT 4.4*) with a membrane of appropriate MWCO to concentrate sought-for proteins while keeping them cautiously in solution.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

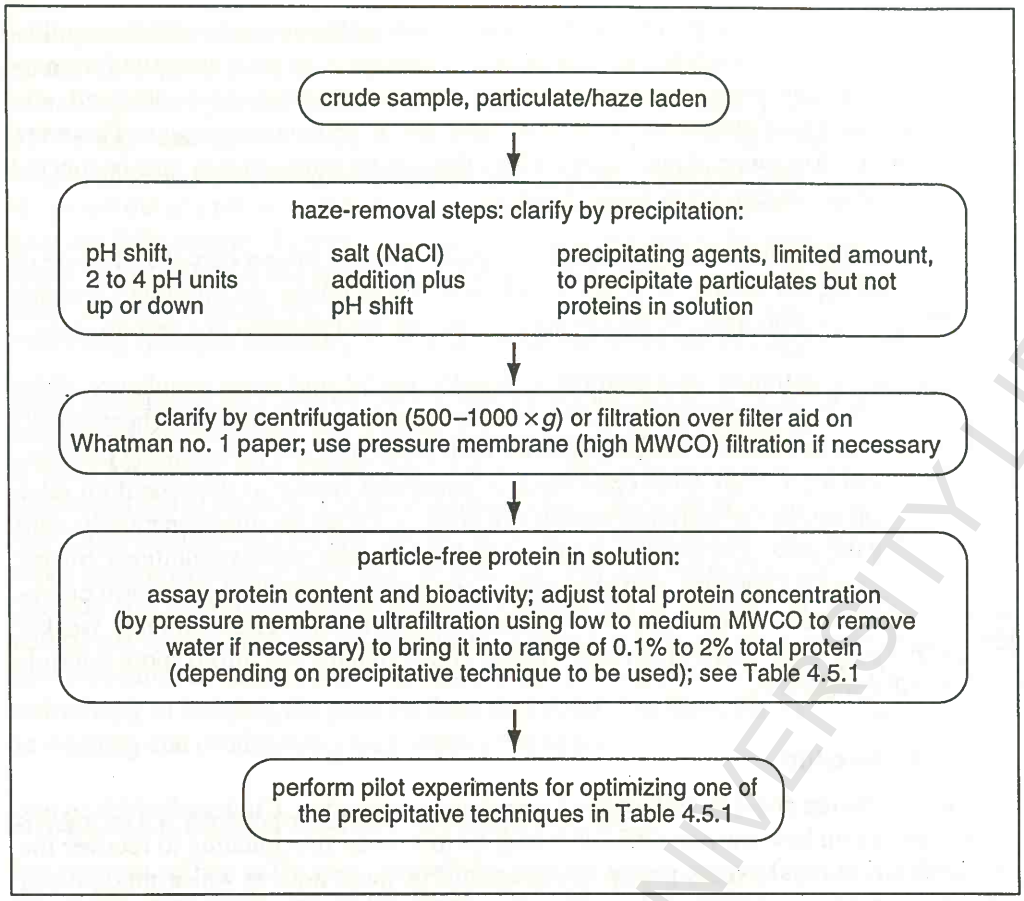

crude sample, particulate/haze laden
haze-removal steps: clarify by precipitation:
pH shift, 2 to 4 pH units up or down
salt (NaCl) addition plus pH shift
precipitating agents, limited amount, to precipitate particulates but not proteins in solution
clarify by centrifugation (500–1000 × g) or filtration over filter aid on Whatman no. 1 paper; use pressure membrane (high MWCO) filtration if necessary
particle-free protein in solution:
assay protein content and bioactivity; adjust total protein concentration (by pressure membrane ultrafiltration using low to medium MWCO to remove water if necessary) to bring it into range of 0.1% to 2% total protein (depending on precipitative technique to be used); see Table 4.5.1
perform pilot experiments for optimizing one of the precipitative techniques in Table 4.5.1

Figure 4.5.1 First steps and options for clearing particulates—e.g., haze and cell walls—from crudes before precipitating proteins. Limited amounts of precipitating agents may be effective in removal of haze from proteins. However precipitating agents for haze removal should not reach lower threshold concentrations for starting precipitation of sought-for proteins. Abbreviation: MWCO, molecular-weight cutoff.

Often it is best not to overconcentrate—i.e., not to attempt to remove nearly all of the water but to get the sought-for proteins to >1% to 2% weight concentration—before precipitating. Unwanted materials such as cell debris, polysaccharides, glycoproteins, and lipoproteins, may precipitate out early, coprecipitate desired proteins, or otherwise become difficult to deal with when they are overconcentrated. In some instances it may be the best choice to dilute samples with a buffer that can also shift pH, then carry out the concentration step. A general starting range of desired protein concentration (clear of haze) is 0.01% to 2%.

Pilot Experiments

After the protein sample has been clarified and concentrated (if needed), aliquots of haze-free sample should be placed in small test tubes or 2-ml plastic centrifuge tubes. A series (Fig. 4.5.3) or a 3 × 3 or a 4 × 4 grid of tubes (Fig. 4.5.11) are used. At first, buffers should be added to the aliquots of sample in rather large steps, at intervals of 1 or even 2 pH units. As the procedure is developed, much narrower pH intervals will be used to zero in on the optimum pH, depending on where bioactivity and protein-specific activity assays locate sought-for proteins in each tube. Likewise, in the first pilot experiments one may make rather large, well-spaced strides in concentrations of precipitating agents added—perhaps 5- to 10-fold differences between pilot samples to locate the neighborhood of precipitation versus nonprecipitation for chosen agents. Then, in a second series, more narrowly spaced conditions should be used to get smoother control over precipitation.

DUKE / CENTRAL UNIVERSITY LIBRARY

Protein precipitation and coprecipitation frequently are quite cooperative (in the equilibrium sense of the term) like other types of phase changes—i.e., protein precipitations may tend to be all-or-none processes. Therefore good control sometimes can be obtained only within small ranges of conditions, such as in variation of precipitating agents. However, the neighborhood or range of parameters where that can be achieved may best be located by the first, wide-ranging set of experiments.

Pilot experiments should make it possible to decide whether the process shall be one of first precipitating out desired proteins, as opposed to first taking out unwanted proteins and perhaps other materials and then bringing down desired proteins in a second stage.

Investigation of optimum temperatures is usually useful and even mandatory. Pilot samples may be kept in refrigerators, at room temperature, and at still higher (heat-shock) temperatures, to find the best temperatures for both protecting and precipitating proteins. Protein precipitation, even when optimized, is sometimes nearly as dependent on rates (kinetics) as on equilibria (thermodynamics). In practice, therefore, attention must be paid to controlling the time of incubation of pilot samples under the various conditions. Strong precipitating agents (synthetic polyelectrolytes are leading examples) may form coprecipitates within a few seconds, but be irreversible. Intermediate and even quite weaker precipitants may take hours to develop protein coprecipitates but afford more controllability and reversibility.

Choice of Procedure

Table 4.5.1 lists ten possible techniques for protein precipitation. Choice of which to use likely depends on how easy or difficult it may be to reverse precipitation to recover the protein(s) and discard excess precipitating agent. For most uses to which precipitated proteins are put—e.g., nutrition, therapy, or simply for assay—the proteins must be freed of precipitating agents. The quite powerful agents, such as synthetic polyelectrolytes, can be difficult or slow (or both) to remove and release from desired proteins. Consequently, synthetic polyelectrolyte coprecipitation may best be used for first bringing down unwanted proteins (i.e., where those precipitates are to be discarded). Remaining proteins in the supernatant may then be captured using a less aggressive but more controllable, reversible precipitating agent from the list in Table 4.5.1 (e.g., Basic Protocols 1 to 5 or 8 to 10).

Removal of Nucleic Acids

Crude proteins from cell cytosols, especially from bacteria and yeasts, may be laden with nucleic acids (RNA and DNA). Nucleic acids may interfere with protein isolation and contribute to haze and colloidal character. It is useful to remove nucleic acids in the early stages of crude workup. Techniques for removal of nucleic acids also help remove some kinds of particulates when the particulates have much anionic character—e.g., many cellular and subcellular membranes. There are two general methods for removing nucleic acids—i.e., precipitation using polyamines (*UNIT 1.3*) and using positively charged anion-exchanger resins (*UNIT 6.1*). Nucleic acids are polyanions that are charged negatively and rather densely with phosphate groups. Naturally occurring "protamines" (Felix, 1960; Budavari, 1989), which are proteins themselves, but very cationic proteins, thoroughly coprecipitate most nucleic acids and ribosomes. However, overshooting the endpoint—i.e., adding more than enough protamine proteins to remove the nucleic acids—may leave excess protamines in samples, contaminating the sought-for proteins and perhaps interfering with their isolation downstream. Accordingly, it may be necessary to determine the endpoint for stopping protamine addition using a suitable spectroscopic technique for nucleic acid content (see *UNIT 7.2*). Synthetic polyethyleneimine (PEI; Sigma) and syn-

BCU IAS CENTRAL UNIVERSITY LIBRARY

thetic polyamino butane compounds (spermine; Sigma) also precipitate nucleic acids, and are useful for large samples (milliliter to liter in volume). A detailed review of PEI (also called Polymin by some vendors) for nucleic acid precipitation is found in Jendrisak (1987). Between 20 and 80 µl of 5% PEI, added per ml of crudes containing 4% (dry weight) nucleic acids, nearly completely removes nucleic acids. This procedure should be carried out at a pH of ~7.5 to 8, where PEI amine groups are fully cationic and nucleic acids are fully anionic. However, PEI is a very able precipitant for proteins with isoionic points below pH 7. If the sought-for protein coprecipitates with both PEI and nucleic acids, as some enzymes do, pH variation and salt extraction (e.g., with NaCl; Jendrisak, 1987) then becomes necessary to free such enzymes from the PEI agglomerate.

Anion exchangers may also be used to remove nucleic acids and have an advantage and a disadvantage compared to polyamines used in solution. The advantage is that an excess of solid exchanger may adsorb less of the sought-for proteins than an excess of such reagents as PEI in solution. The disadvantage is that—because solid resins can only adsorb very large macromolecules like nucleic acids at the resin surface (not in its interior)—rather large quantities of resin may be required. The Amberlite IRA-400 series of resins (Sigma) and Dowex-1 anion exchangers (Sigma) are adsorbants for nucleic acids in neutral pH ranges. The DEAE-celluloses (Sigma), which are quaternized amines and therefore cationic, also are capable of adsorbing nucleic acids to remove them from crudes preliminary to isolating the proteins from the crudes. See Basic Protocol 2 for directions on washing and conditioning such resins prior to use.

SELECTIVE PRECIPITATION BY SALTING OUT

BASIC PROTOCOL 1

Ammonium sulfate is the best, first-choice salt for initial development of a salting-out program to precipitate sought-for proteins (see Fig. 4.5.2). Other salts may be chosen, and frequently they yield successful results. However because of sulfate's kosmotropic and protein-molecule exclusionary powers (see Commentary), ammonium sulfate is most preferred and best understood. Sulfate salting out may be both thermodynamically limited and rate-limited. If precipitation of proteins does not occur quickly after addition of estimated amounts of sulfate salt, it may be best to wait for a period, perhaps a few hours, to allow time for precipitates to form. This becomes necessary for dilute protein solutions—less than ~1% in overall concentration. For several mammalian enzymes, the threshold for protein precipitation and solubility—i.e., the visibly seen development of precipitation versus proteins remaining dissolved (apparent equilibrium)—occurs at ~2 to 5 mg/ml protein in 2 to 3 M ammonium sulfate, at neutral pH (Scopes, 1987).

The pH and temperature also are likely to be important. If the isoelectric pH of the sought-for protein is known, it may be used as the center point around which to vary pH, in increments of ~0.5 to 1 pH unit, in initial searches for optimum conditions. A number of proteins, peptide antibiotics, and drugs are best salted out as coprecipitates with SO_4^{2-} as a bound counterion. Hence their optimal pH for salting out may be acidic with respect to their isoelectric point. Figure 4.5.2 illustrates the main steps in salting out, starting with mixtures of proteins in solution without particulate matter.

In both this protocol and Alternate Protocol 1, the sample is first carried through all steps on a small scale for a pilot experiment to determine optimal conditions. The procedure is then scaled up to purify the protein sample, using those optimal conditions. Figure 4.5.3 illustrates the procedures for such a pilot experiment.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

Materials

Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
Appropriate pH buffer
Ammonium sulfate

Additional reagents and equipment for dialysis (*UNIT 4.4 & APPENDIX 3B*)

1. Dialyze protein samples against a pH buffer or a pH buffer/ammonium sulfate mixture having a sulfate concentration below that needed to start precipitation (see *UNIT 4.4 & APPENDIX 3B*).

   By carrying out dialysis, three objectives may be achieved at the expense of an added step. First, it is possible to dialyze out unwanted lower-molecular-weight materials—e.g.,

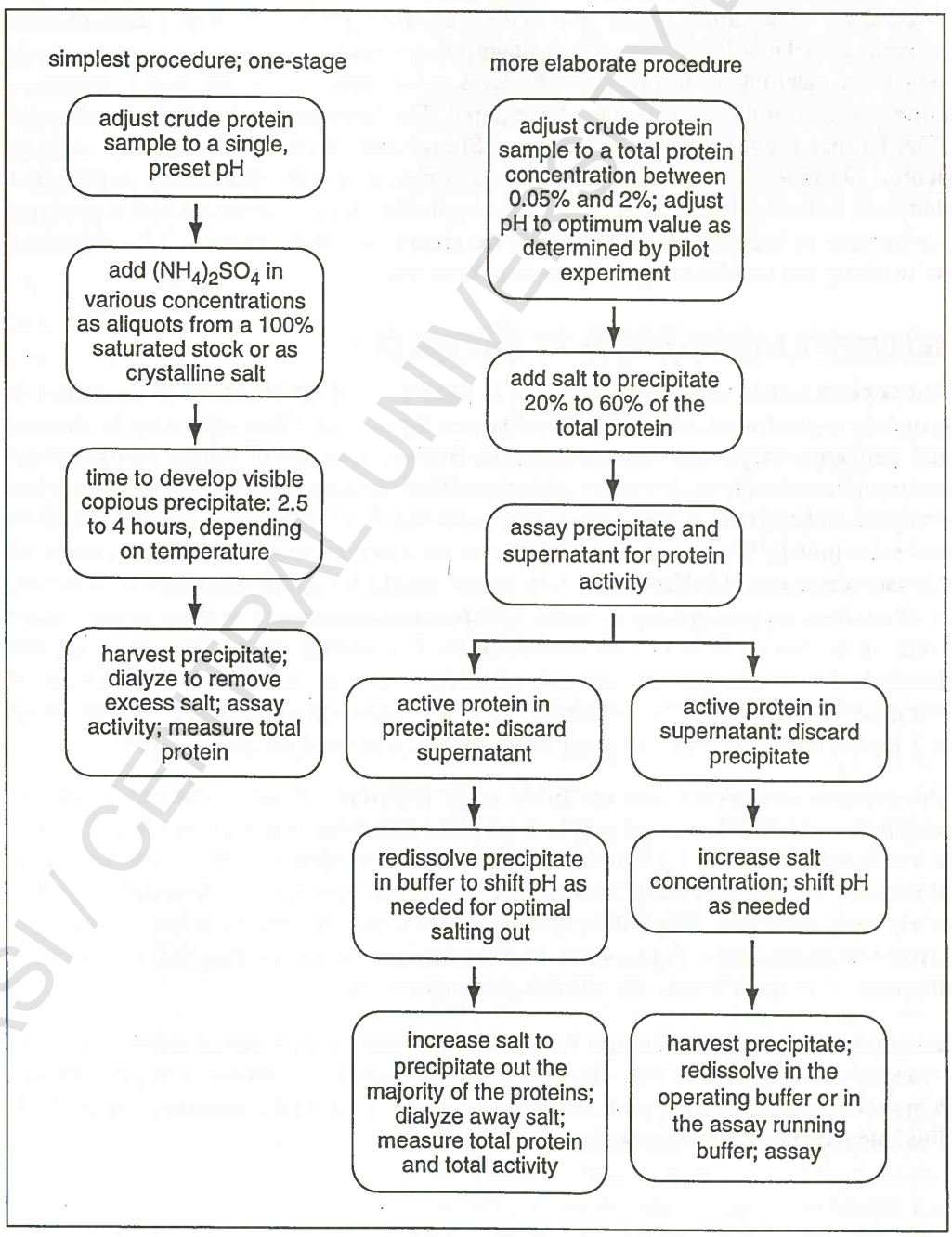

simplest procedure; one-stage
adjust crude protein sample to a single, preset pH
add $(NH_4)_2SO_4$ in various concentrations as aliquots from a 100% saturated stock or as crystalline salt
time to develop visible copious precipitate: 2.5 to 4 hours, depending on temperature
harvest precipitate; dialyze to remove excess salt; assay activity; measure total protein
more elaborate procedure
adjust crude protein sample to a total protein concentration between 0.05% and 2%; adjust pH to optimum value as determined by pilot experiment
add salt to precipitate 20% to 60% of the total protein
assay precipitate and supernatant for protein activity
active protein in precipitate: discard supernatant
redissolve precipitate in buffer to shift pH as needed for optimal salting out
increase salt to precipitate out the majority of the proteins; dialyze away salt; measure total protein and total activity
active protein in supernatant: discard precipitate
increase salt concentration; shift pH as needed
harvest precipitate; redissolve in the operating buffer or in the assay running buffer; assay

Figure 4.5.2 Two-salting out protocols, single-stage and two-stage, starting from a particle-free aqueous crude preparation (e.g., cell or tissue extract or centrifuged fermentation broth).

RCU IASI / CENTRAL UNIVERSITY LIBRARY

sugars, amino acids, and components of fermentation broth or cell extract. Second, dialysis makes it possible to bring the protein sample smoothly to the prescribed pH. Finally, dialysis allows the investigator to adjust the protein sample to a known beginning salt concentration—that of the dialyzing buffer. Such dialysis may be carried out in dialysis bags (cellulose acetate tubing), or by pressure membrane concentrators with membranes having nominal molecular-weight cutoff points below that of the sought-for protein.

2. Set up a pilot experiment (see discussion of Pilot Experiments in Strategic Planning) to determine the optimal protein concentration, pH, salt concentration, temperature, and incubation time.

   See Critical Parameters and Troubleshooting for guidelines on how to vary these parameters. See Figure 4.5.3 for illustration of a small-scale pilot experiment.

   pH adjustments generally are performed by addition of acid (HCl), base (NaOH, KOH, or ammonia), or a buffer. At the same time, it is usually necessary to avoid adding too much overall volume, so the buffering capacity of the sample itself must be taken into account. If feasible, it is generally best to adjust pH of proteins with ~0.1 or 0.2 M HCl, or comparable concentrations of aqueous ammonia. Trying to move the pH appreciably with buffers, if dilute buffers must be used, increases overall volumes and decreases salt concentration. That, in turn, may require the use of considerably more salt to reach a specified concentration, in contrast to operations where pH has been adjusted by use of HCl or NaOH.

   CAUTION: *The protein solution should be stirred well during addition of strong acid or strong base. Beware of changes in temperature.*

3. Add ammonium sulfate to the optimal concentration (see Critical Parameters and Troubleshooting). Incubate (for the optimal period of time) until a precipitate forms.

   The salt may generally be added as a saturated solution. However, there is an advantage to adding the salt in solid form, as increases in volume are thereby lessened. There are few indications that addition of solid ammonium sulfate does harm to proteins already in solution.

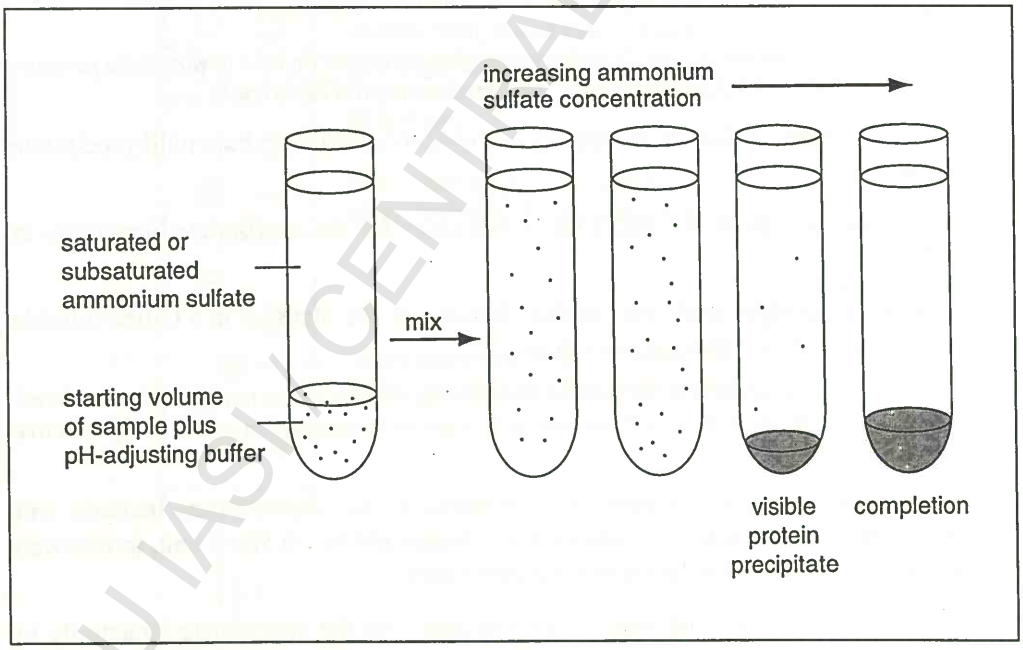


Figure 4.5.3 Salting out—small-scale investigation. Maximum ammonium sulfate concentration ~4 M. Development of precipitate depends on the protein and unwanted materials which sometimes precipitate out in samples or fractions, apart from the sought-for protein. Other adjustable parameters are pH, temperature, and concentration of starting sample.

EMU IAS CENTRAL UNIVERSITY LIBRARY

4. Collect the precipitate by simple decantation (pouring off the supernatant).

   Ammonium sulfate precipitates are usually dense and settle out readily. However, if simple settling and decantation will not work, a slow-speed centrifugation (10 to 100 × g) should easily suffice.

5. Redissolve the precipitate in a buffer suitable for the next step (e.g., running buffer for electrophoresis or assay buffer).

   High concentrations of K^+ and other salts are not compatible with SDS-PAGE and isoelectric focusing. K^+ salts of SDS are poorly soluble. Large amounts of ammonium ion grossly interfere with some total protein assays (UNIT 3.4). Therefore, excess salt should be dialyzed away by simple cellulose-tubing dialysis (UNIT 4.4 & APPENDIX 3B) or by pressure ultrafiltration (UNIT 4.4).

6. Perform a bioassay for the protein of interest.

   If the sample is to be diluted considerably for the bioassay, the redissolved ammonium sulfate may not need to be dialyzed beforehand. However, if appreciable salt interferes (as in SDS-PAGE and total protein assays), the assay may require a dialyzed sample input.

ALTERNATE PROTOCOL 1

SELECTIVE PRECIPITATION BY STEPWISE SALTING OUT

A stepwise salting-out procedure is used to fractionate a crude mixture into two portions, only one of which retains the desired protein. The result is a partial purification of the desired protein from a crude mixture of proteins; this may improve the extent of recovery in subsequent purifications.

In this protocol, as in Basic Protocol 1, the sample is first carried through all steps on a small scale for a pilot experiment to determine optimal conditions. The procedure is then scaled up to purify the protein sample, using those optimal conditions. Figure 4.5.3 illustrates the procedures for such a pilot experiment.

1. Adjust the concentration and pH of the crude protein solution to the optimal values as determined by a pilot experiment (see discussion of Pilot Experiments in Strategic Planning).

   See Critical Parameters and Troubleshooting for guidelines on how to vary these parameters. See Figure 4.5.3 for illustration of a small-scale pilot experiment.

2. Add ammonium sulfate at an appropriate concentration. Incubate until precipitate forms.

3. Collect the precipitate and supernatant and assay for the appropriate bioactivity in each.

4a. *If the active fraction is the precipitate:* Redissolve the material in a buffer suitable for assay or the next general procedure.

   If the next steps are not to be performed immediately, the precipitate need not be redissolved. It should be stored with its salt because salts, especially ammonium sulfate, are protective and stabilizing.

4b. *If the active fraction (or part of it) remains in the supernatant:* Increase salt concentrations by 5% to 10% and perhaps change pH by ~0.5 to 1 unit, in renewed pilot experiments and in larger-scale precipitation.

5. Collect the precipitate and supernatant and assay for the appropriate bioactivity in each.

6. Repeat steps 4b and 5 until the protein of interest is obtained in as clean a preparation as possible.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

SELECTIVE PRECIPITATION BY ISOIONIC PRECIPITATION: COLUMN METHOD

Proteins frequently are least soluble and most precipitable when they are isoionic. In the isoionic, salt-free state, protein molecules are in their most compact, least hydrated conformation—a phenomenon that is closely related to the condition of proteins at their isoelectric point. The distinction between isoionic and isoelectric properties is drawn in detail by Tanford (1961). Deionization using a column (this protocol) or dialysis (see Alternate Protocol 2) aim at rendering proteins isoionic to precipitate them. Two important parameters determine solubility of many proteins: solution pH with respect to each protein's isoionic point (pI) and the low salt concentration (zero to 0.1 to 0.2 M salt). A number of proteins—e.g., β-lactoglobulin—are sharply dependent on these parameters. Accordingly, if these parameters are well controlled and carefully adjusted, the solubility/precipitability behavior of a protein will be a practical basis for scaleable, selective isolation of the desired protein. The column method used in this protocol is appropriate only for proteins that remain soluble at their isoionic point.

In addition to adjusting proteins to their isoionic pH, the general method described here, using mixed-bed resin deionization, is able to strip away all salts from proteins. Salts, even in small concentrations (<0.05 M in many cases) often have large effects on protein solubilities and therefore on precipitability. Inorganic salts tend to "salt in" many proteins, thus enhancing their solubilities. Accordingly, sharp control and thorough removal of salts as desired is an important part of the general technique of isoionic precipitation.

Figure 4.5.4 shows a Dintzis deionization column with a lower, main mixed bed of anion-cation exchange resin beads in their respective OH^- and H^+ form. Two categories

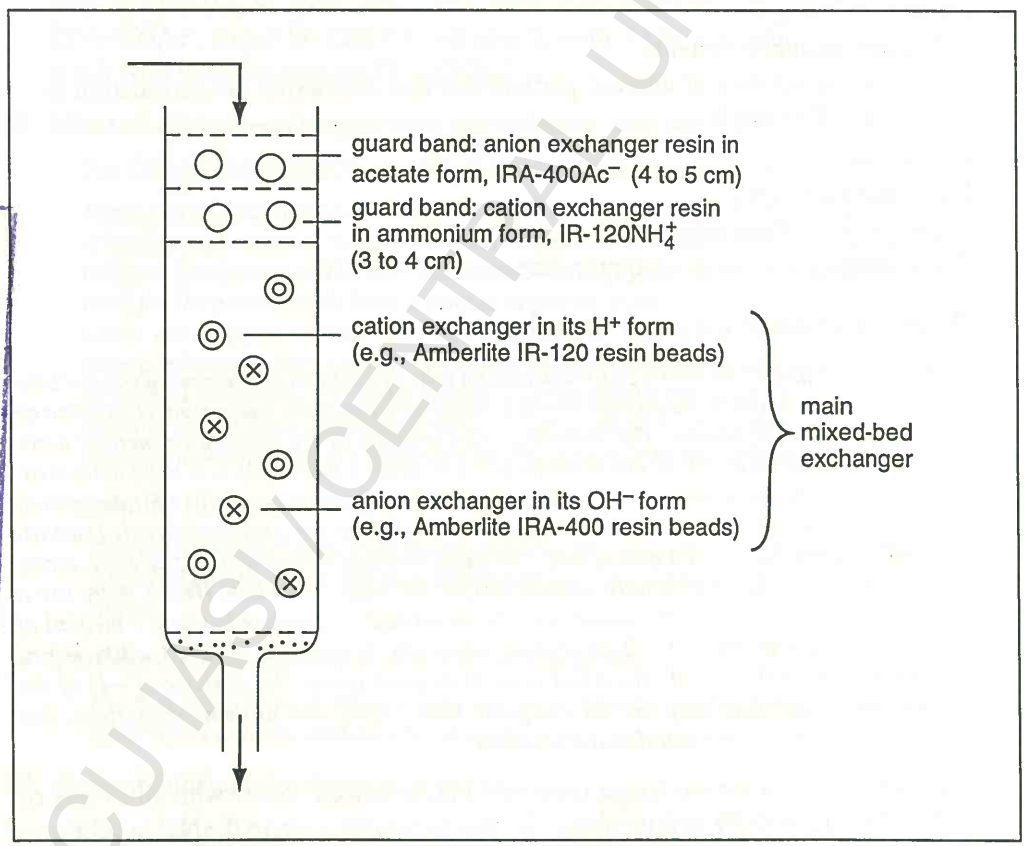

guard band: anion exchanger resin in acetate form, IRA-400Ac⁻ (4 to 5 cm)
guard band: cation exchanger resin in ammonium form, IR-120NH₄⁺ (3 to 4 cm)
cation exchanger in its H⁺ form (e.g., Amberlite IR-120 resin beads)
anion exchanger in its OH⁻ form (e.g., Amberlite IRA-400 resin beads)
main mixed-bed exchanger

Figure 4.5.4 Deionization column, Dintzis design, described in Edsall and Wyman (1958). Passage through this column renders the protein salt-free and automatically adjusted to the protein's isoionic pH. This column is used for proteins that remain soluble at their isoionic point.

B. C. U. „M. Eminescu" IAŞI CHIMIE

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

of salt ions are removed by the mixed bed: (1) the supporting electrolyte—i.e., the free salt in the solution; and (2) salt ions that are counterions, which are in effect bound to the protein if the protein is not isoionic. If the pH of the starting solution is below the isoionic point, the protein is a cation with anion counterions, commonly Cl^-. If the pH is greater than the isoionic point, the protein is a net macroanion with cationic counterions, commonly K^+ or Na^+. In the mixed bed, cationic and anionic salt ions are exchanged for H^+ and OH^-, respectively, going on to simply form water. Protein molecules deprived of all mobile counterions are quantitatively forced to the single pH at which they can exist without mobile counterions or protein-bound ions—i.e., the isoionic pH. Therefore the effluent of the column is isoionic protein at the isoionic pH, in plain water. The net salt exchange and entrapment processes are similar to conventional desalting by resin ion exchangers, except the effluent is buffered by the isoionic protein and adjusted to its pH.

In the original Dintzis column, illustrated in Figure 4.5.4, guard exchanger bands may be positioned at the top of the column above the main bed. They exchange incoming salt cations and anions (or protein counterions) for NH_4^+ ion and acetate ion from the resins. This is intended to protect incoming proteins from direct mixing with the equivalent of strong acid and strong base, respectively, from the mixed-bed strong exchange resins in their respective H^+ and OH^- form. The NH_4^+ and acetate ions are exchanged and removed in the main bed, and the isoionic protein sample flows through.

Materials

Cation exchanger: Amberlite IR-120H^+ resin (Sigma or Aldrich)
Anion exchanger: Amberlite IRA-400Cl^- resin (Sigma or Aldrich)
~0.5 and 1 M NaOH or KOH
0.5 to 1 M HCl
0.2 M acetic acid
0.2 M ammonium hydroxide
Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)

Plastic beakers
Large Buchner funnel
Whatman no. 1 filter paper
Chromatography column of appropriate size

1. Determine exchange capacity required.

   Necessary capacity for deionizing and trapping salts depends on the number of counterions and therefore on pH of the protein solution, concentration of the supporting electrolyte or buffer, and overall volume. The following is an example of the calculation method used. Assume that 50 ml of 5% serum albumin in 0.2 M acetate buffer, pH 4, is to be deionized. On the basis of the titration curve (with a pI of 5.0 for albumin; see UNIT 8.2 for discussion of titration curves), ~35 counterions (e.g., Cl^-) are present per protein molecule (mol.wt. 67,000), amounting to ~1.3 meq of ions. The buffer contains ~10 meq of total acetate, along with 10 meq of Na^+, which were counterions for the buffer's acetate. Hence, there are a total of 22 meq of ions to be exchanged. The mixed bed is calculated to hold 1 meq/ml of settled mixed-bed resin, so ~12 ml of resin, minimally, is required. Thus a 2 × 30–cm bed including 80 to 90 cm^3 of mixed-bed resin is in good excess (by a factor of ~4) of the minimum required to strip out all inorganic ions, buffer, and protein counterions. See Critical Parameters for additional discussion.

2. Place IR-120H^+ cation-exchange resin in a plastic beaker, cover with an excess of ~0.5 M NaOH or KOH, and incubate ~20 min to convert resin to the Na^+ or K^+ form. Transfer to a large Buchner funnel containing Whatman no. 1 filter paper, wash with water, then transfer to a beaker and cover with an excess of 0.5 to 1 M HCl. Incubate

BCU IASI CENTRAL UNIVERSITY LIBRARY

20 min, then transfer to a Buchner funnel containing filter paper and wash with water again. Repeat this exchange cycle at least one more time before mixing the exchange resins.

This full-cycle treatment removes undesirable low-molecular-weight hydrolysis fragments from resin polymers, which otherwise would bind strongly to many proteins and elute with them when proteins are passed through the column. Each step requires ~20 min, to allow penetration, diffusion, and exchange of ions with the interiors of the resin beads. More detailed specifications are given in full treatises describing exchangers, such as Kressman (1957) and in the manufacturers' literature.

3. Place IRA-400Cl^- resin in a plastic beaker, cover with an excess of ~1 M NaOH or KOH, and incubate ~20 min to convert resin to the OH^- form. Transfer to a large Buchner funnel containing Whatman no. 1 filter paper and wash with water to remove all excess OH^- ion.

   The resin is now IRA-400OH^-. It should be kept cold and should not be exposed to air during long periods of storage because the CO_2 in air slowly neutralizes the IRA-400OH^- to a bicarbonate form. The resin should be taken through this full exchange cycle within a few days prior to use.

   Anion exchangers, both old and new, which are made of quaternary amine polymers, self-hydrolyze and slough off organic amines, especially when stored in their hydroxyl form (e.g., IRA-400OH^-). Accordingly, IRA-400 in whatever starting form (Cl^- or OH^-), needs be taken through a full exchange cycle. Each step requires ~20 min, to allow penetration, diffusion, and exchange of ions with the interiors of the resin beads. More detailed specifications are given in full treatises describing exchangers, such as Kressman (1957) and in the manufacturers' literature.

4. Prepare resin for preexchange guard bands as follows. React IRA-400OH^- (from step 3) with 0.2 M acetic acid, then wash with water to prepare anion-exchanger IRA-400Ac^-. React IR-120H^+ (from step 2) with 0.2 M ammonium hydroxide, then wash with water to prepare IR-120NH_4^+.

5. Mix and load ion-exchange resins into columns (also see Fig. 4.5.4).

   See Critical Parameters for guidelines on preparing the mixed-bed resin.

   Some precautions need be taken to prevent unmixing—i.e., formation of separated bands of cation and anion exchangers in their respective H^+ and OH^- forms—when filling the column. The density of IR-120H^+ is appreciably larger than that of IRA-400OH^-. The resins used for the guard bands have a similar disparity in density. If large volumes of premixed resins are dropped through several centimeters of water, they unmix on the way down. Hence, mixed-bed resins should be dropped in small increments, with forward flow of water, to prevent unmixing. The same general considerations pertaining to procedures such as packing and loading chromatographic columns, dead-volume minimization, and prevention of channeling that are described in UNIT 8.4 apply to this column technique.

6. Determine flow rate required.

   If there is adequate capacity in the mixed bed to trap all salts, a 2 × 30–cm column deionizes 50 to 100 ml of a 1% to 5% protein solution in ~3 hr. Suitable flow rates average 0.5 to 1 ml/min.

7. Perform column separation on crude protein solution.

   The first cut, ~20 to 30 ml of eluate, will be quite dilute as the protein solution displaces water in the column, and may be set aside.

8. Perform bioassay for the protein of interest.

9. Recycle used resins.

   Mixed-bed resins previously used for deionization—e.g., of serum proteins—are good general binding agents and can be reused. After being used for deionizing a protein, the

BOUVIER CENTRAL UNIVERSITY LIBRARY

resin should be flushed with 0.2 M HCl to remove any protein and to prevent bacterial growth. Mixed-bed resin pairs of the IRA-400 and IR-120 series, as well many of the Dowex resin pairs, separate well from one another in saturated KCl. Crystalline KCl can be added to a slurry of used mixed-bed resin in water. As KCl saturation approaches, the anion exchanger that is least dense—i.e., the Cl^- form—floats upward. The cation exchanger in K^+ form sinks. The anion-exchange resin slurry is decanted off the top, away from the cation exchanger, isolating both resins. These may then be individually recycled back to their respective OH^- and H^+ forms, whereupon they may again be mixed as needed (see steps 2 and 3).

ALTERNATE PROTOCOL 2

SELECTIVE PRECIPITATION BY ISOIONIC PRECIPITATION: DIALYSIS METHOD

One of the older means of rendering proteins salt-free or nearly salt-free (i.e., isoionic) is dialysis (also see *UNIT 4.4 & APPENDIX 3B*). However, two problems frequently arise with conventional dialysis: (1) when appreciable amounts of protein are present, osmotic effects result in swelling of dialysis bags as salt diffuses outward; and (2) often it is quite uncertain where the isoionic point is, even if it is feasible to deionize by dialysis against a buffer. The resin deionization method (see Basic Protocol 2), sometimes called Dintzis' method (Edsall and Wyman, 1958), automatically adjusts a protein precisely to its isoionic pH without prior knowledge of it.

Proteins that precipitate immediately when deionized are troublesome in the column method described in Basic Protocol 2 and Figure 4.5.4. Precipitates plug the column and pose the problem of separating precipitated protein from resin beads. β-lactoglobulin and a number of other proteins behave in this way; they have solubilities very sensitive to low salt conditions near the isoionic pH. This protocol (also see Fig. 4.5.5) describes a means of overcoming this problem. Proteins to be deionized and precipitated in isoionic form are confined in dialysis tubing with a molecular weight cutoff that will retain the protein of interest and the mixed-bed resin is placed outside as a slurry of 40 to 60 g of mixed-bed resin (50 to 80 ml of wet resin; see Basic Protocol 2) per 200 to 400 ml of dialyzing solvent. Salt ions and protein counterions exchange through the membrane and are trapped outside in the exchanger resins. After exchange is completed, precipitated protein in the dialysis tubing is recovered by centrifugation of the bag's contents. This general technique requires several hours. Salting out and diffusion slow as the protein concentration inside the dialysis bag decreases; hence this method is slower than the flow-through column method (see Basic Protocol 2).

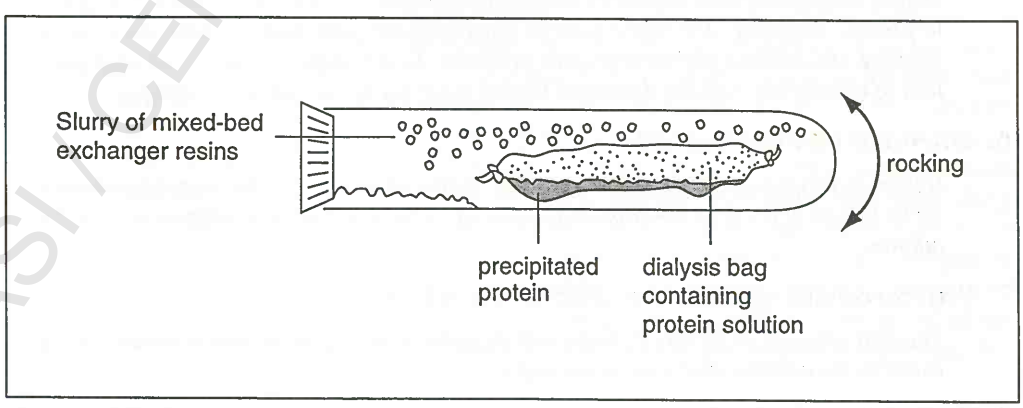


Figure 4.5.5 Deionization by dialysis. When proteins are insoluble and precipitate at their isoionic point, deionization is accomplished by placing dialysis tubing containing the protein sample in a slurry of mixed-bed resin exchangers (i.e., a mixture of IRA-400OH^- and IR-120H^+) and incubating with rocking. Proteins insoluble at their isoionic point precipitate inside the bag. The mixed-bed exchange resins remove free salts, forcing the protein to its isoionic pH. This method is considerably slower than the column method.

RCU IASI / CENTRAL UNIVERSITY LIBRARY

SELECTIVE PRECIPITATION USING A TWO-CARBON (C_2) ORGANIC COSOLVENT

Protein precipitation using very cold ethanol or acetone is a method that was developed over a century ago. Both of these cosolvents are still frequently used. Ethanol surpassed acetone for precipitating proteins at the time E. J. Cohn, and co-workers at Harvard perfected the Cohn fractionation (Cohn et al., 1946) for large-scale preparation of plasma proteins—albumin, fibrinogen, prothrombin, and γ-globulins. The Cohn procedures are still employed, although modified. One of the advantages ethanol has over acetone pertains to fire hazard. Ethanol is less volatile than acetone and generates less flammable vapor under otherwise similar conditions. Both cosolvents, when handled in in >0.1-liter amounts, should be used in ventilated rooms.

Materials

Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
Buffering system
Ethanol
Acetone
Centrifuge

1. Adjust the crude protein solution of interest to the selected pH.

   If the operating buffer (used as the solvent in the separation) or pH-calibration buffers (used to check pH-meter calibration) are to be selected for low-temperature operation, the tables by Alner et al. (1967) and Perrin and Dempsey (1974) give formulations for several buffers together with their pH values at temperatures in the 0°C range.

2. Arrange for good stirring and good metering to control mixing of organic cosolvents with the aqueous proteins.

   It is necessary to suppress local high concentrations of C_2 cosolvents. It is also necessary to suppress the heat generated by mixing of cosolvents, especially of ethanol with water, to avoid undue denaturation and temperature increases greater than ~2°C above the set temperature. Dilution of organic cosolvents with water or with aqueous running buffer before they are added to the protein solution may repress generation of heat during mixing and may also repress formation of large local concentrations of cosolvents at the point at which they are mixed with aqueous protein. However excessive dilution of organic cosolvents may add to the final volumes necessary to reach precipitation thresholds. A 3% to 8% (v/v) volume of water in organic cosolvent is a suitable dilution for achieving protein precipitation before overall volumes become excessive.

3. Cool the two components—i.e., protein in aqueous buffer and cosolvent system to be used—to the running temperature. Do not allow temperatures to rise more than 2°C above the running temperature.

   Lower temperatures can be obtained by using an auxiliary refrigeration unit pumping coolant through an exchanger coil or by incubating C_2 cosolvents in the deep freeze. Salt/ice baths are capable of reaching approximately –10°C.

4. Slowly mix the the cosolvent system into the aqueous protein solution while monitoring the mixture with a thermometer.

   If the temperature rises much during cosolvent (or cosolvent/buffer) addition, slow the rate of solvent addition down to let the cooling system remove excess heat. Predilution of organic cosolvents (see annotation to step 2) may be helpful here.

5. Collect precipitate by centrifugation or allow precipitate to settle out. Remove supernatant and redissolve precipitate by adding water or suitable aqueous buffer.

BCIT IRSI / GERRARD UNIVERSITY LIBRARY

Organic cosolvents admixed with aqueous systems usually decrease the density of the solvent mixture, which may aid settling or centrifugation of precipitated protein. However, the centrifuge should also be cooled down to the precipitation running temperature. Some centrifuges become surprisingly warm, depending on the vintage and specific features. A test of the actual centrifugation temperature is advisable; this is done simply by immersion of a thermometer in centrifuge vessels after a run.

6. Perform a bioassay for the protein of interest.

BASIC PROTOCOL 4

SELECTIVE PRECIPITATION USING C_4 AND C_5 ORGANIC COSOLVENTS

Figure 4.5.6 outlines the technique for three-phase partitioning (TPP) using *t*-butanol in conjunction with ammonium sulfate (Lovrien, et al., 1987). Some C_5 and C_6 cosolvents—e.g., alcohols such as pentanol—may be useful in some cases. However C_5 and larger cosolvents tend to be so insoluble in water as to be of limited use. The cosolvent used here, *t*-butanol, is infinitely soluble in water alone, but in the presence of large concentrations of ammonium sulfate it forms a second phase. Simultaneously, proteins are precipitated, forming a third phase intermediate between the lower (aqueous) and upper (organic) liquid phases (see Fig. 4.5.6). The various phases are mutually saturated with respect to one another. The protein precipitate is usually well developed and easily retrieved by low-speed centrifugation.

TPP behavior depends on protein molecular charge (precipitated products are often sulfate salts of protein). Therefore, the pH of the aqueous phase is an important parameter for modulating protein precipitation, partitioning, and extraction of unwanted compounds. When proteins have been extracted from their original source using detergents, TPP or related methods such as Morton's *n*-butanol extraction technique (Morton, 1950) are frequently effective in removing detergents, lipids, and pigments from the sought-for protein.

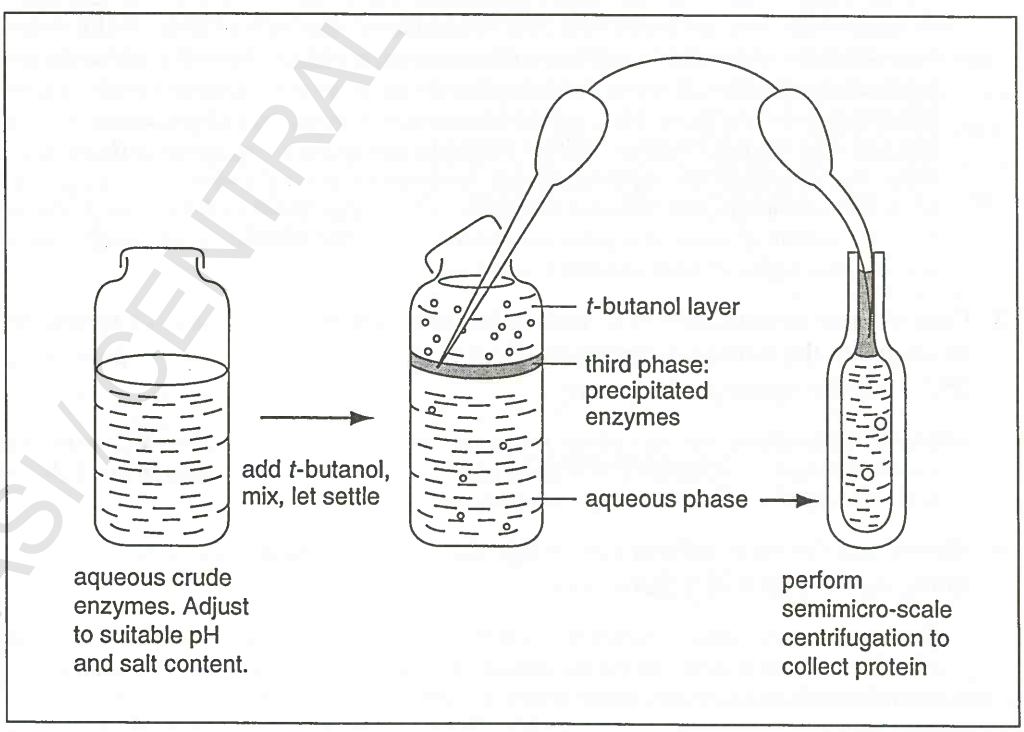

t-butanol layer
third phase: precipitated enzymes
aqueous phase
add t-butanol, mix, let settle
aqueous crude enzymes. Adjust to suitable pH and salt content.
perform semimicro-scale centrifugation to collect protein

Figure 4.5.6 Procedure for precipitation of proteins from aqueous solutions using *t*-butanol. This figure applies to both large-scale precipitation and small-scale pilot experiments. The *t*-butanol layer develops when adequate salt is present in the aqueous solution. Both the *t*-butanol layer and the aqueous lower phase act as extraction solvents as well as partitioning systems.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Materials

Ammonium sulfate
Buffering system
Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
C_4 organic cosolvent (e.g., *t*-butanol, analytical grade)

1. Add ammonium sulfate and buffer to adjust the salt concentration and pH of the crude protein solution of interest appropriately.

   The amount of ammonium sulfate used is determined by pilot experiments and is in general less than that used for simple salting out (20% to 100% saturation); i.e., the amount added is less than sufficient to induce precipitation.

2. Add ~0.5 to 1 ml C_4 organic cosolvent per ml crude protein solution, mix, and incubate until three phases form.

   If only technical grade t-butanol is available it may be either recrystallized (the melting point of pure t-butanol = 25.5°C) at room temperature, or distilled (the boiling point = 82°C).

3. Transfer the middle phase containing the precipitated protein to a centrifuge tube of volume approximately twice that of the protein precipitate. Centrifuge at $10 \times g$, 4° to 25°C.

   See Critical Parameters for discussion of the separation of phases.

4. Redissolve the precipitate in an appropriate buffer and perform suitable bioassay.

SELECTIVE PRECIPITATION USING PROTEIN EXCLUSION AND CROWDING AGENTS AND OSMOLYTES

BASIC PROTOCOL 5

Protein exclusion and crowding agents and osmolytes are neutral, very water-soluble compounds—such as sugars, organic diols, and polymers—used in concentrations of 2% to 20% to precipitate proteins. They operate primarily by entropically driven "crowding" (Herzfeld, 1996), by preferential hydration (Timasheff, 1992), or a combination of both mechanisms. Such agents push proteins out of solutions in the mechanical/physical sense and in the thermodynamic sense. At concentrations greater than ~5% by weight, exclusion agents and osmolytes occupy so much volume on the molecular scale that they leave less room in solution available for protein molecules.

Crowding and molecular-exclusion polymers that have been used for protein precipitation include polyethylene glycol (PEG), polypropylene glycol (PPG), polyvinyl alcohol (PVA), methylcellulose, dextran, and hydroxypropyldextran. Osmolytes and zwitterions used for protein precipitation include sucrose, 2-methyl-2,4-pentanediol (MPD), raffinose, maltose, sarcosine, and betaine. These compounds are readily available from suppliers such as Sigma and are able to protect and precipitate a number of proteins. Development of each protein-precipitation program requires attention to pH, temperature, and other factors that modulate parameters governing protein solubility and precipitability. The literature is scant concerning rate limitation of protein precipitation by these parameters. It is likely that in lower concentration ranges of proteins, neutral polymers, and osmolytes (<1% to 2%), precipitation rates may be slowed. Accordingly the general method should be treated as a rates (kinetics) problem when precipitates are not quickly formed.

The sample is first carried through all steps on a small scale for a pilot experiment to determine optimal conditions. The procedure is then scaled up to purify the protein sample, using those optimal conditions. Figure 4.5.3 illustrates the setup for such a pilot experiment.

BOU MACHINE UNIVERSITY LIBRARY

Materials

Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
Appropriate pH buffer
Precipitating agent: neutral polymer or osmolyte

1. Set up a pilot experiment (see discussion of Pilot Experiments in Strategic Planning) to determine the optimal concentrations of protein and precipitating agent as well as the optimal pH, temperature, and incubation time.

   In preliminary searches, use a series of test tubes as in Figure 4.5.3. On addition of components to each other, mixing or strong vortexing may be needed because some neutral polymers are highly viscous.

2. Prepare protein solution and solution of precipitating agent.

   Neutral polymers should be dissolved in water before adding them to aqueous proteins. Moderate heating may speed dissolving. It is poor practice to attempt addition of dry polymers directly to dissolved proteins as the polymers tend to form gummy masses that dissolve very slowly.

   Mechanisms here depend on molecular crowding; hence concentrations of proteins and dissolved polymers to be added must be maximized. These components dilute each other when mixed. If protein samples need pH adjustment before the mixing step, use buffers in fairly concentrated form—or even 0.2 to 1 M acid or base if feasible—to keep volumes minimized and component concentrations maximized.

4. Mix protein solution and solution of precipitating agent. Incubate (for the optimal period of time) until a precipitate forms.

   If mixtures of proteins are very crude and only one or two proteins are sought after, it may be preferable not to drive the precipitation too hard. In such cases it may be best to accept more slowly developed precipitates to maximize the amount of unwanted protein and excess precipitant left behind in the supernatant and the amount of sought-after protein captured in the precipitate.

5. Collect the precipitate.

   Trapped and coprecipitated neutral polymers in precipitates may need be "reversed" or split away later to release proteins free of polymers. How this can be done depends on the chemistry of the systems, and perhaps also on the chromatographic properties of the precipitated proteins. In the case of dextrans, modest amounts can be degraded by dextranase enzyme. Coprecipitated proteins may be captured on ion-exchange matrices, whereupon neutral polymer may be eluted and the protein displaced from the exchanger using a pH shift and/or salt (see UNIT 8.2).

6. Perform bioassay for the protein of interest.

BASIC PROTOCOL 6 SELECTIVE PRECIPITATION USING SYNTHETIC AND SEMISYNTHETIC POLYELECTROLYTES

Four categories of polyelectrolytes can precipitate proteins. First, there are the synthetic polyelectrolytes—e.g., vinyl polymers such as polyacrylate (PAA) and polymethacrylate (PMA); acid salts (polyanions); and polyethyleneimine (PEI, a polycation). Next, there are the semisynthetic polyelectrolytes—carboxymethylcellulose (CMC), sulfated cellulose, and sulfated dextrans. The third category includes the protamines (very basic proteins), and the fourth includes naturally occurring sulfate polysaccharides—e.g., heparin and chondroitin sulfate. Heparin and related biochemical polyelectrolytes are not generally used on a large scale because of their cost; they are used in research for precipitating narrow classes of proteins such as β-lipoproteins. For a recent, short review

PCU IAS/CENTRAL UNIVERSITY LIBRARY

Table 4.5.2 Synthetic Polyelectrolyte Coprecipitating Agents for Proteins

Polyelectrolyte	Proteins preciptated	References
Polyacrylate	Lysozyme, hemoglobin	Sternberg and Hershberger (1974)
Polyacrylate	Eight industrial enzymes	Sternberg (1976)
Polyacrylate	Ferrihemoglobin, catalase, bovine serum albumin	Berdick and Morawetz (1954)
Polymethacrylate	Cellulase enzymes	Whitaker (1953)
Polymethacrylate	Serum albumin, hemoglobin	Morawetz and Hughes (1952)
Polyethyleneimine	Recombinant (RecA) protein	Shibata et al. (1981)
Polyethyleneimine	RNA polymerase II	Jendrisak and Burgess (1975)
Polyethyleneimine	β-galactosidase fusion protein	Niederauer et al. (1994)
Polyethyleneimine	Restriction endonucleases	Bickle et al. (1977)
Carboxymethylcellulose (CMC)	β-lactoglobulin, α-lactalbumin	Hidalgo and Hansen (1971)
Polylysine	Fibrinogen	Sela and Katchalski (1959)
Copolymer Eudragit	Protein A	Mattiasson and Kaul (1994)
Tri-block polyampholytes	Trypsin, RNase A, lysozyme	Patrickios et al. (1994)
Polydimethyldiallylammonium chloride	Serum albumin	Dubin et al. (1987)

of polyelectrolytes for protein precipitation see Singh (1995). Table 4.5.2 lists a number of polyelectrolytes in each category and cites references regarding their use. Figure 4.5.7 outlines the main steps for polyelectrolyte-mediated protein precipitation.

Frequently, the main practical problem with this technique is in reversing coprecipitation—i.e., dissolving the coprecipitate in a way that releases proteins and splits away the polyelectrolyte. In the case of polyacrylate and polymethacrylate, these polymers bind calcium and barium ions rather strongly. Thus protein/polyacrylate coprecipitates may first be shifted upward in pH, which renders proteins and polyelectrolytes negatively charged and helps them redissolve. After this, Ca^{2+} or Ba^{2+} is introduced to precipitate away the polyelectrolyte and release the protein. Several papers describing examples of the use of this technique are referred to in the Commentary (see Critical Parameters and Table 4.5.2).

The general range of starting protein concentrations from which proteins may be coprecipitated in fair to good yield (at optimum pH) is ~0.05% to 3% (w/v) protein. The range of quantities of polyelectrolyte required, given a suitable kind of polyelectrolyte, tends to be less than the amount of protein. By weight, ~2% to 20% polyelectrolyte with respect to the protein is generally required. The variation between systems is large, partly because polyelectrolytes have large variations in charge density—e.g., completely ionized polyacrylic acid carries one anionic charge per 56 g polymer; carboxymethylcellulose, substituted on each glucose unit, carries one anion per 240 g polymer if completely ionized.

Materials

Water-soluble polyelectrolytes (Aldrich or Sigma)
Crude protein solution of interest

Test tubes
Glass or plastic cuvettes
Spectrophotometer

BCU IAS CENTRAL UNIVERSITY LIBRARY

1. Choose a polyelectrolyte.

For proteins not previously investigated, first experiment with from two to four easily available, water-soluble polyelectrolytes. Polyacids (e.g., polyacrylic acids), and polybases (e.g., polyethyleneimine), can be titrated with NaOH, or HCl respectively, to two or three pH units, respectively, above or below the target protein's isoelectric point (if known). Such polyelectrolytes are buffers themselves, as are proteins, over the pH range of acid/base titration. Hence, it is unnecessary to use additional conventional inorganic buffers to control pH. If polyelectrolytes and proteins are mismatched in pH before mixing, the mixture may shift in pH; the pH should be measured, if necessary, to determine H^+ transfer on mixing the main components.

2. Prepare stock polyelectrolyte solutions.

Initial stock polyelectrolyte solutions should be prepared in water in concentrations three to ten times the final coprecipitating concentrations. Dilutions will be performed in the succeeding steps.

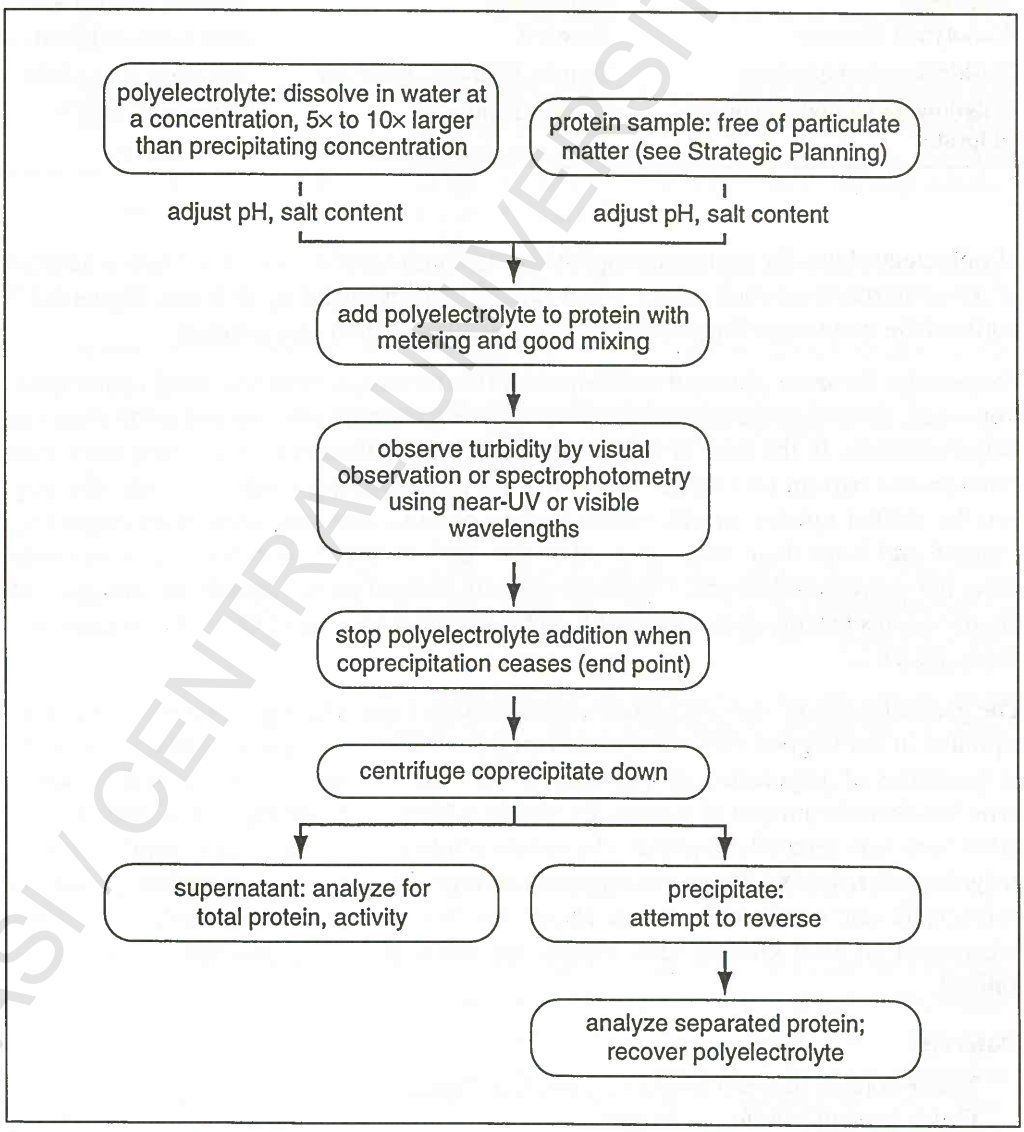

polyelectrolyte: dissolve in water at a concentration, 5× to 10× larger than precipitating concentration
protein sample: free of particulate matter (see Strategic Planning)
adjust pH, salt content
adjust pH, salt content
add polyelectrolyte to protein with metering and good mixing
observe turbidity by visual observation or spectrophotometry using near-UV or visible wavelengths
stop polyelectrolyte addition when coprecipitation ceases (end point)
centrifuge coprecipitate down
supernatant: analyze for total protein, activity
precipitate: attempt to reverse
analyze separated protein; recover polyelectrolyte

Figure 4.5.7 Principal steps in protein coprecipitation with polyelectrolytes. "Metered" addition is slow addition (with good stirring) of aliquots of polyelectrolyte stock to the protein sample near the threshold of precipitate formation as seen by eye or spectrophotometric turbidity determination.

RCU IASI / CENTRAL UNIVERSITY LIBRARY

3. Render protein sample clear and nonparticulate (see discussion of Clarification in Strategic Planning).

4. Set up several test tubes on a small scale, adding ~2 ml sample and 0.05 to 1 ml of each polyelectrolyte stock solution at three or four different pH values to construct a matrix of polyelectrolyte/protein mixtures. Observe the point at which turbidity appears.

   The intent is to roughly locate (by visual comparison to detect turbidity) the coprecipitation starting point, the coprecipitation end point, the relative concentrations of polyelectrolyte and protein to be used, and the optimum pH. This enables construction of a rough phase diagram (i.e., a plot of turbidity versus variables such as relative amount of polyelectrolyte).

   Also see discussion of Pilot Experiments in Strategic Planning.

5. Set up a more finely grained pilot experiment as in the previous step, but carry out turbidity measurements in glass or plastic cuvettes. For scale-up, use controlled rates of polymer addition (i.e., adding polymer to the protein sample slowly in reproducible aliquots of stock solution near the threshold of precipitate formation), temperature control (water bath), and mixing using a magnetic stir-bar. Measure spectrophotometric absorbance in the 400- to 500-nm range and determine where overt coprecipitation begins.

   Sharply climbing absorbances mark the ranges where turbidity forms as a result of particulate coprecipitates. Coprecipitation is in large part a rate process as well as an equilibrium process. If the aim is to achieve selection of particular proteins, rather than precipitation of all proteins, the best strategy may to bring down some protein(s) in a first-stage separation of the coprecipitate, followed by a second stage in which polyelectrolyte is again added.

6. Carry out full-scale separation. Centrifuge to collect coprecipitate; retain the supernatant and analyze for total protein activity.

7. Reverse coprecipitation to free captured protein from synthetic polymers.

   This step should start with resuspension of particles in water or in buffers to shift pH to a position where polyelectrolyte and protein have similar charges—e.g., both anionic or both cationic. Separating out polyelectrolytes is often achieved by addition of metal ions (Ca^{2+} or Ba^{2+} in the form of chlorides) to carboxylate polyelectrolytes. An attempt may also be made to adsorb polyelectrolytes on bead resins of opposite charge—e.g, on DEAE cellulose for carboxylated polyelectrolytes.

   In developing a procedure for polyelectrolyte removal and release of sought-for proteins, measurement of supernatant protein concentration and/or activity if the protein is an enzyme should facilitate monitoring of protein release. Most of the polyelectrolytes in Table 4.5.2 have low A_{280} absorbances (and small molar extinction coefficients). Hence simple measurement of A_{280} should locate the conditions at which proteins are released into solution, provided there are no particulates present. However, such particulates cause very strong scattering and turbidity in the UV range, far beyond the absorbance of equivalent amounts of truly dissolved protein.

8. Perform bioassay for protein of interest and recover polyelectrolyte.

BCU IAS CENTRAL UNIVERSITY LIBRARY

BASIC PROTOCOL 7

SELECTIVE PRECIPITATION USING METALLIC AND POLYPHENOLIC HETEROPOLYANIONS

Under strongly acid conditions, inorganic and some organic strong anions—known as heteropolyanions—ably precipitate proteins. Inorganic anions used under acid conditions for this purpose include perchlorate, tungstate, phosphotungstate, molybdate, phosphomolybdate, tungstosilicate, and ferrocyanide. Inorganic ions (used under acid to neutral conditions) include sulfosalicylate, picrate, and diverse plant polyphenols and tannates (also see Critical Parameters).

Prior to overt precipitation, protein molecules under strongly acidic conditions are acid-expanded and remain soluble. When the heteropolyanions are introduced and allowed to bind, the protein molecules are forced back to a compact, poorly hydrated conformation. Molecular expansion and contraction of the protein in solution, which leads to precipitate formation, may be followed by biophysical tools (Fink, 1995). Driven further with additional heteropolyanions, such as perchlorate and tungstate in ~0.2% to 2% or 3% concentration, proteins coprecipitate with these anions in dense aggregates that easily settle or centrifuge down.

In some procedures for precipitating proteins, perchloric acid ($HClO_4$) is the precipitating agent of choice because it is ultraviolet-transparent above 250 nm. Perchloric acid precipitates whole protein molecules, but not their low-molecular-weight fragments—amino acids and oligopeptides. This is the basis for analysis with many proteolytic enzymes—at the end of the proteolytic cleavage assay, whole or intact proteins are precipitated out by a few percent $HClO_4$. Proteolytic fragments are left in the supernatant for subsequent A_{280} measurement, giving a rather direct measure of the amount of hydrolysis that occurred before adding the $HClO_4$.

BASIC PROTOCOL 8

SELECTIVE PRECIPITATION USING HYDROPHOBIC ION PAIRING (HIP) ENTANGLEMENT LIGANDS

Hydrophobic ion pairing (HIP) is coprecipitation of proteins using flexible hydrocarbon "tails" of alkane anions (detergents) that have their anion head groups bound in the target proteins. Figure 4.5.8 illustrates the composition of such a coprecipitate. Organic anions, especially sulfonates and sulfates, often bind strongly to proteins bearing cationic sites.

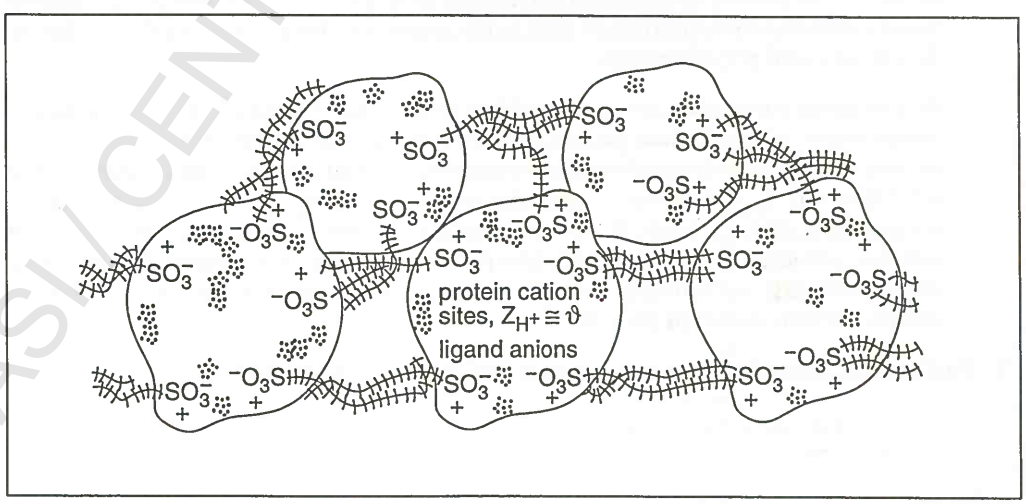


Figure 4.5.8 HIP coprecipitate. Strong anions (sulfate or sulfonate-bearing C_{10}–C_{14} alkane tails, e.g., dodecyl sulfate) bind to protein cationic sites forming ligand-protein complexes. Complexes draw together via tail-tail hydrophobic entanglement, aggregate, and coprecipitate. At maximum coprecipitating efficiency with ~10^{-5} M protein, the stoichiometric ratio of bound ligand anions (ν) is to cation side chains (Z_{H^+}) in the protein molecule is close to 1:1.

RCU IASI / CENTRAL UNIVERSITY LIBRARY

The organic alkane groups entangle and bind to each other in a manner similar to tail/tail association in micelles. Entanglement draws protein/ligand complexes together as shown in Figure 4.5.8. Such complexes go on to aggregate, then coprecipitate. Ordinary detergents such as dodecyl sulfate, (DDS^-) in correct protein/detergent ratios are able protein precipitants. Such ligands, especially detergent anions, need to be carefully regulated with regard to concentration. Suitable overall concentrations of precipitating detergents are usually 10^{-4} to 10^{-5} M when used with overall protein concentrations ranging from 0.01% to 0.001% (w/w) or 10^{-4} to 10^{-6} M protein. If suitable ligand concentrations are exceeded, the ligands go on to form micelles and solubilize proteins rather than precipitate them. Inside their optimum coprecipitation range with respect to total protein concentration, these ligands are rather generally effective towards many proteins.

Protein coprecipitation by HIP is frequently easy to reverse. The precipitate is redissolved in solvent with a pH ~2 or 3 units more alkaline than the coprecipitation pH. Detergent anions may be quantitatively stripped out with strong ion exchange resins, usually in the chloride form—e.g., Amberlite IRA-400Cl^- or Dowex-1Cl^-, or equivalent resins. Released proteins—e.g., enzymes in the supernatant—are then ready for assay.

SELECTIVE PRECIPITATION BY MATRIX-STACKING LIGAND COPRECIPITATION

BASIC PROTOCOL 9

Matrix-stacking ligands function (see Fig. 4.5.9) to draw protein molecules together in dense complexes that aggregate and coprecipitate protein. Ligand/ligand association forms a matrix host; protein molecules trapped in the matrix are guests. Occasionally, these ligands cocrystallize proteins and peptides (Conroy and Lovrien, 1992). Ligand organic tails bear rather rigid coplanar groups that stack and bind to one another π-face–π-face fashion, in a manner similar to stacking of aromatic bases in nucleic acid double helices. Tail-to-tail matrix association is reinforced by short alkane groups on the periphery of the tails—i.e., methyl, ethyl, and *t*-butyl substituents. The alkyl substituents lend hydrophobicity and promote water displacement, which helps drive tail-to-tail association. The structures of a few anionic ligands, are shown in Figure 4.5.10.

As with flexible-entanglement ligands (see Basic Protocol 8), the more rigid matrix-stacking ligands achieve maximal coprecipitation efficiency when ν ligands are bound per protein molecule with close to a 1:1 ratio of ν to Z_{H^+} —i.e., $\nu = \sim Z_{H^+}$, where Z_{H^+} = protein net cationic charge from H^+ titration of side chains. Hence coprecipitation depends on pH and how pH governs the protein charge, Z_{H^+}. Stacking and entanglement ligands in their most useful overall concentrations of 10^{-4} to 10^{-5} M often protect proteins as well as precipitate them. Such ligands tighten protein conformation, sometimes profoundly (Matulis and Lovrien, 1996). Protection often starts in solution before overt coprecipitation. The capacity to protect proteins against several kinds of degradation is an advantage of matrix-stacking ligand coprecipitation.

Two variables around which to design early protein coprecipitate experiments are illustrated in Figure 4.5.11. The following steps may be used for developing a matrix-ligand coprecipitation protocol.

Materials

Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
Potential ligands (Fig. 4.5.10; also see Aldrich catalog)
Appropriate buffers
Dowex-1Cl^- resin
Small test tubes or 2- to 3-ml plastic microcentrifuge tubes

BOULAS/CENTRAL CONNECTICUT STATE UNIVERSITY LIBRARY

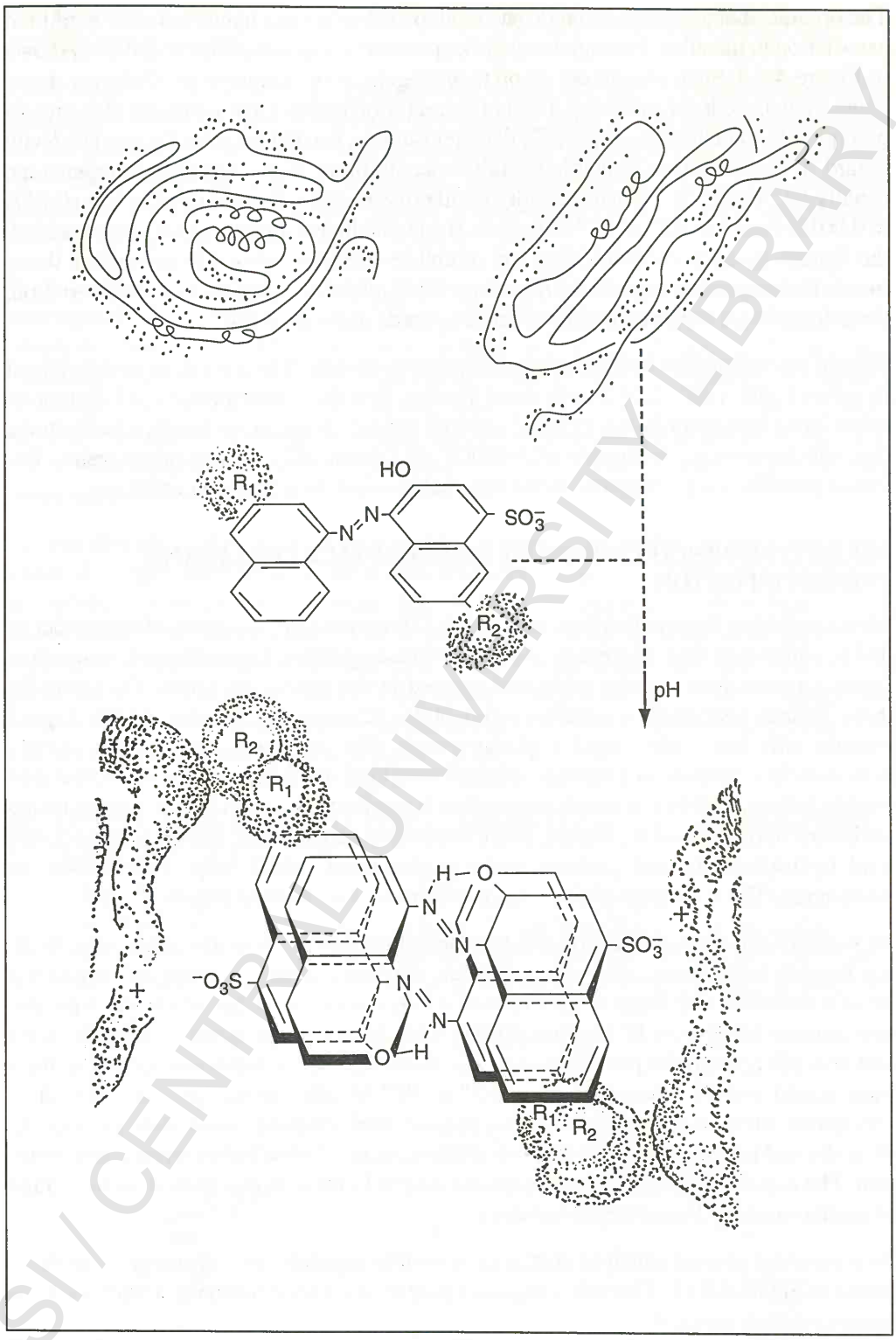


Figure 4.5.9 Molecular structure-function basis for matrix-stacking ligand coprecipitation of proteins. Protein molecules initially soluble in water and bearing a positive net charge attract strong anion (sulfonate) ligands. Such ligands associate via organic tail/organic tail stacking interaction, reinforced by alkane R-groups. Ligand/protein complexes are thus drawn together and coprecipitate out of solution.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

1. Set up grids of small test tubes or 2- to 3-ml plastic microcentrifuge tubes as in Figure 4.5.11. In each grid, vary two parameters versus one another to optimize coprecipitation—e.g., pH versus ligand concentration (with a fixed total protein concentration), pH versus kind of ligand, and other pairs of variables as described in Figure 4.5.11.

2. If possible, determine protein concentrations in samples to be precipitated (see *UNIT 3.4*).

   Usually if stacking and entanglement ligands are to function, proteins need be adjusted in pH in such a way that proteins of 20- to 60-kDa molecular weight carry a Z_{H^+} of ~15 to 40 net cationic charges. If protein molecular weights are unknown, assume ~40,000 Da, or assume the molecular weights typical of the kinds of proteins being worked with (e.g., common microbial and digestive proteases range from 20 to 35 kDa, and coagulation proteases from 50 to 75 kDa).

3. Determine isoionic point for the protein and optimal pH for precipitation.

   Expect to adjust the pH of some samples to between 2.5 and 5 if information from titration curves or amino acid analysis is not available for predicting Z_{H^+}. Preferably this is done

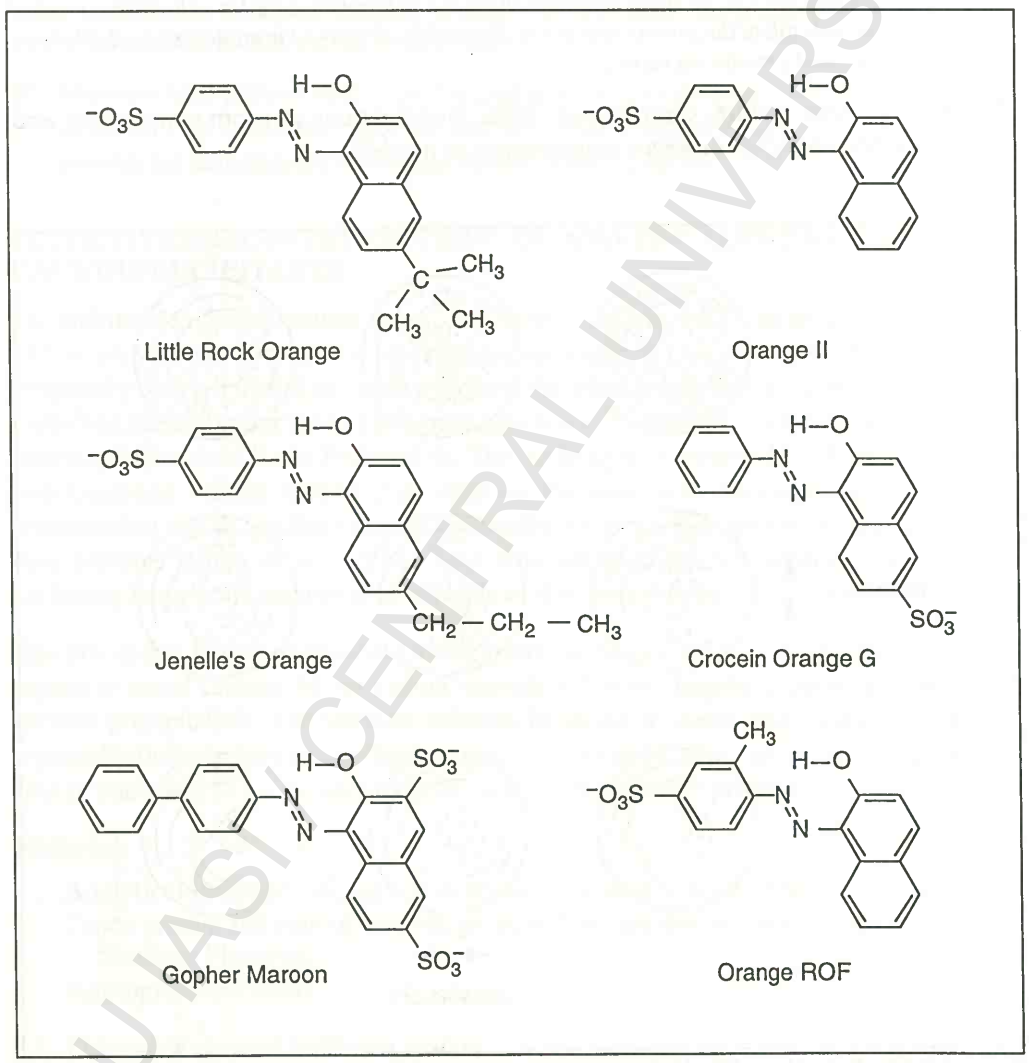


Figure 4.5.10 Anionic ligands synthesized with reinforcing alkane groups (Little Rock Orange and Jenelle's Orange) are very strong protein coprecipitants. Commercially available dye anions—Orange II, Crocein Orange G, and Orange ROF (available from Aldrich)—frequently are also efficient coprecipitating ligands for cationic proteins in pH ranges 2 to 4 units below the isoionic pH of the protein.

RU IASI CENTRAL UNIVERSITY LIBRARY

by direct acidification with ~0.01 M HCl accompanied by pH monitoring. Alternatively, buffers—e.g., glycine or formate buffers—are recommended. Do not adjust protein pH into acid regions before adding the ligands unless there is a need for protection by ligands in control experiments. If sought-for proteins are fragile at acid pH, ligands should be added before titrating the proteins to the lowered pH. Control (contrast) experiments may use the opposite sequence—i.e., acidify first and add ligand second. If there turns out to be much difference in the results obtained via the two sequences, protection of proteins from acid-pH denaturation is indicated.

4. Experiment with two or three ligands from Figure 4.5.10 or choose other potential ligands from among the azoaromatic sulfonate dyes listed in the Aldrich catalog. Prepare ligand stock solutions.

   Ligands are preferably dissolved in water or in dilute buffer at the running pH. Stay below the limit of ligand solubility, which can vary considerably, in making stock solutions. Upper solubility limits for a number of such ligands are 0.01 to 1 mM.

5. Add ligands from ligand stock solutions to protein sample in various volume ratios—e.g., ~1 vol stock solution to 2 vol protein solution and vice versa. Briefly vortex the mixtures.

   Extreme ratios—e.g., 10 vol of ligand solution to 1 vol protein samples—should be avoided as these will dilute the protein too much. If possible, arrange for modest mutual dilutions of ligand and proteins on mixing.

6. Optimize temperature, carrying out initial investigations at room temperature and proceeding to higher or lower temperatures as needed.

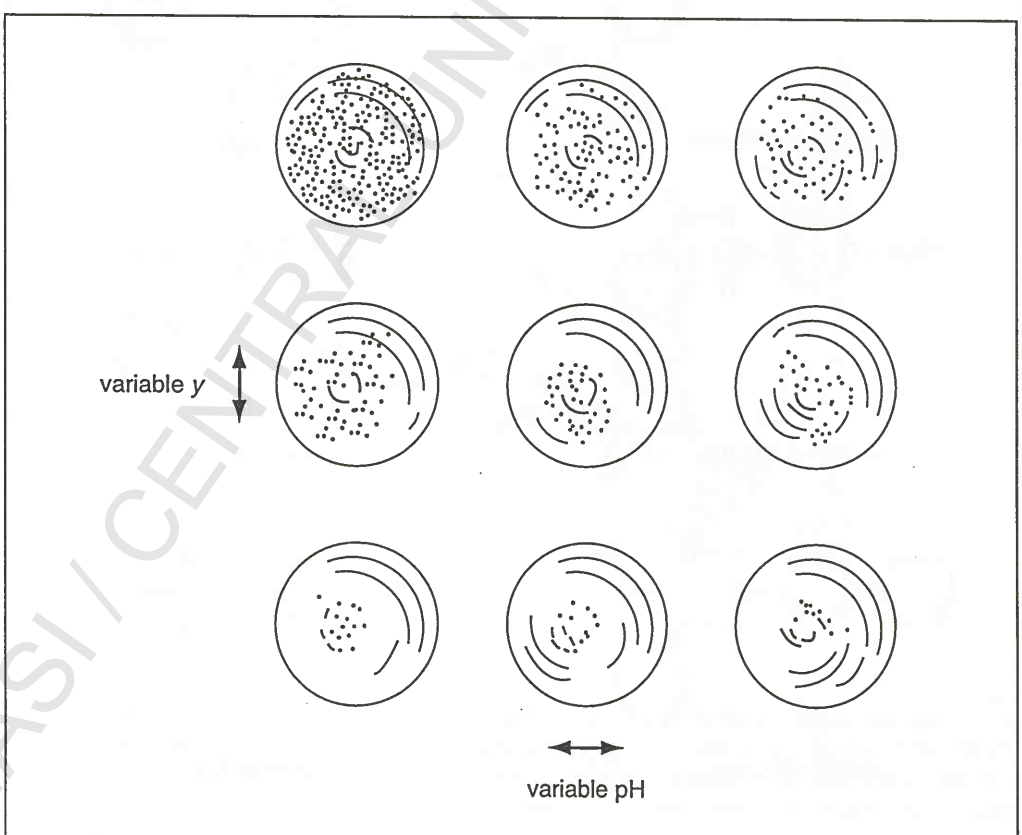


Figure 4.5.11 Sets of small test tubes or plastic microcentrifuge tubes, 2 to 3 ml in volume, may be used to to explore optimum conditions for matrix-ligand/protein mutual coprecipitation by varying pairs of important parameters. Five principal variables determining coprecipitation are: kind of ligand; ligand-protein ratio (*y*; often between 2 and 20); initial pH; temperature; and concentrations of auxiliary coprecipitating agents such as Zn^{2+} (10^{-3} to 10^{-4} M).

PCU IASI / CENTRAL UNIVERSITY LIBRARY

7. Determine rate of precipitation giving the optimal yield.

Matrix-ligand coprecipitation is sometimes rate-limited when protein concentrations start from <0.1% (w/v). If this is the case, incubate ligand/protein mixtures for a few hours. Rather often, coprecipitation occurs rapidly; however, to avoid including unwanted compounds in coprecipitates, do not drive the coprecipitation reaction too rapidly (if possible).

8. Collect coprecipitates by centrifugation at 5 to 50 × *g*. Perform bioassay of the coprecipitates and, if necessary, of the supernatant for the protein of interest.

In some systems, e.g., peroxidase isolations from crudes, myriad unwanted proteins precipitate out first and the peroxidase enzyme is left in the supernatant.

9. Redissolve coprecipitates by shifting pH upward with 0.1 M acetate, imidazole, or phosphate buffers, pH 5 to 7. Add ~100 mg of Dowex-1Cl$^-$ resin and incubate 5 to 15 min with gentle shaking to trap the anionic ligands. Monitor the completeness of trapping via the color of the matrix ligand.

Approximately 100 mg of resin per milliliter of the higher-pH redissolving buffer provides a good excess of trapping capacity. See Basic Protocol 2 for discussion of the capacities of anion exchange resins. Dowex-1 resins have ~1 meq of capacity per ml (wet volume) of resin beads. Complete trapping is easily followed from matrix ligand's color. Released proteins in solution should be colorless.

10. Measure total protein (and total released activity if an enzymatic or other bioassay is available) to determine specific activity, total recovered activity, and recovered protein for comparison to samples before coprecipitation.

SELECTIVE PRECIPITATION USING DI- AND TRIVALENT METAL CATION PRECIPITANTS

BASIC PROTOCOL 10

Di- and trivalent metal cations—e.g., Zn^{2+}, Mn^{2+}, Ca^{2+}, and Al^{3+} (general symbol M^{2+} or M^{3+})—provide two means for precipitating proteins: (1) they may act directly, or (2) frequently they are useful as auxiliary agents for other precipitation methods such as the crowding method using neutral polymers (see Basic Protocol 5) or matrix-stacking ligand coprecipitation (see Basic Protocol 9). This technique is reviewed by Rothstein (1994) and Gurd and Wilcox (1956). Zinc ions are the ions most commonly used for direct precipitation and as auxiliary agents, especially when working with plasma and coagulation proteins (Cohn et al., 1950). Zinc ions are required for insulin crystallization, including large-scale industrial production of this polypeptide.

The discussion below represents a small sector of the landscape of chemical behavior in regard to metal cations. Metal cations provide a diverse, largely undeveloped means of protein precipitation. The intrinsic solution behavior of metal ions (without proteins), especially their chelation-coordination reactions, must be taken into account for application of such ions to the development of protein precipitation methods.

Materials

Analytical Reagent (AR)-grade metal salts (generally as chloride or nitrate)
Crude protein solution of interest, particle-free (see discussion of Clarification in Strategic Planning)
Appropriate buffer(s)

1. Determine optimal buffering system.

The precipitation system may be buffered in two respects: (1) toward H^+ ions and pH in the conventional way and (2) also with respect to metal ions, using metal ion buffers which are coordinating and chelating agents (Perrin and Dempsey, 1974). A number of com-

BOURSE CENTRAL UNIVERSITY LIBRARY

pounds—e.g., EDTA, citrate, and several amino acids—can serve both buffering purposes. Evidence that compounds serve as one or both of these types of buffer is that pH dependency and/or metal ion concentration dependency of protein precipitation is rather sharply dependent on the kind of buffering compound. Zwitterionic buffers such as HEPES may be useful for controlling pH because they have much less tendency to bind or coordinate with metal ions than conventional Sorenson buffers (e.g., phosphate). Because proteins are also buffers, their self-buffering action may suffice to control pH if present in concentrations ≥0.3%.

2. Select an appropriate metal cation and precipitate protein.

Di- and trivalent metal cations (M^{2+} and M^{3+}) are vigorously precipitated by salts and inorganic buffer anions, which are present in protein extracts and samples at various stages of isolation. Sulfate (from salting out; see Basic Protocol 1 and Alternate Protocol) powerfully precipitates Ca^{2+} as $CaSO_4$. Solubilities of <100 mg/ml or even <10 mg/100 ml are seen with $Zn(NH_4)PO_4$, $Mn_3(PO_4)_2{\cdot}7H_2O$, $MnCO_3$, $Mg(NH_4)PO_4{\cdot}6H_2O$, $MgCO_3$, $CaCO_3$, and a number of calcium phosphates. Sulfates and carbonates of Ba^{2+} are extremely insoluble (Hogness and Johnson, 1954). Before committing a metal ion as a precipitating agent: (1) review general behavior with respect to precipitability by buffer and salt anions in the protein solvent; (2) carry out control experiments with protein-free samples to find if the solvent(s) are likely to produce massive precipitates of simple inorganic compounds with the metal ion; and (3) arrange to remove interfering anions if the anions are in concentrations ≥0.05 M or if concentrations of ions such as $SO_{4^{-2}}$ and $CO_{3^{-2}}$ re <0.05 M.

Trivalent metal ions such as Al^{3+} and Fe^{3+} form a complicated series of oxo-, hydrate, and hydroxylated species in water. They go on to form insoluble hydroxides (solubility products 10^{-30} to 10^{-40}; Hogness and Johnson, 1954), at neutral and even somewhat acidic pH. Metal cations of this kind are interesting protein precipitants (Gurd and Wilcox, 1956). However, instead of complexing with protein molecules, they may form very insoluble hydroxyl gels and flocculate proteins by adsorption on the gels at pH ≥5. Hence development of protein precipitation with such cations may need be to confined to rather acidic pH regions.

It is not necessarily a disadvantage if low-molecular-weight salts or other solvent components precipitate M^{2+} first, before proteins are precipitated. Inorganic precipitates of sulfate, phosphate, and carbonate may form first. These can then be centrifuged away, whereupon more metal ions are added to bring down the sought-for protein, in a second stage of precipitation. Total protein measurements (UNIT 3.4) are used to locate the starting point for protein precipitation by the metal ions.

3. Analyze precipitate.

Analysis of metal content of coprecipitates and their supernatants is useful for characterizing products and following how variations in procedures determine product composition. The most generally accessible methods are spectrophotometric, using color-developing reagents capable of quantitating metal cations at 10^{-4} to 10^{-5} M concentrations (Sandell and Onishi, 1978). Figure 4.5.12 shows a calibration curve for Zn^{2+} analysis using pyridylazoresorcinol (PAR; Aldrich). The optimal pH for use of PAR is ~7 (Cheng et al., 1992) in HEPES or other nonchelating buffer. 0.1 mM PAR is prepared in 0.05 M HEPES buffer, pH 7.0, then dilutions of Zn^{2+} (as zinc acetate) are prepared in the same buffer at various concentrations ranging from 5 to 30 μM. The PAR reagent may be made up as a 1 mM stock solution and diluted 10-fold to provide the working reagent at 0.1 mM. Spectrophotometric samples are made by mixing 1.0 ml of 0.1 mM PAR with 1.0 ml of each of the Zn^{2+} dilutions. A_{495} is then measured versus a reference of equal PAR concentration without zinc. The slope of the plot in Figure 4.5.12 is the molar extinction coefficient, $\varepsilon_{495} = 4.5 \times 10^4\ M^{-1}\ cm^{-1}$. An example of Zn^{2+} release from a zinc enzyme monitored by PAR color development is given in Hunt et al. (1985). Zinc analysis is a criterion for the identification and purity of insulin.

PCU IASI CENTRE UNIVERSITY LIBRARY

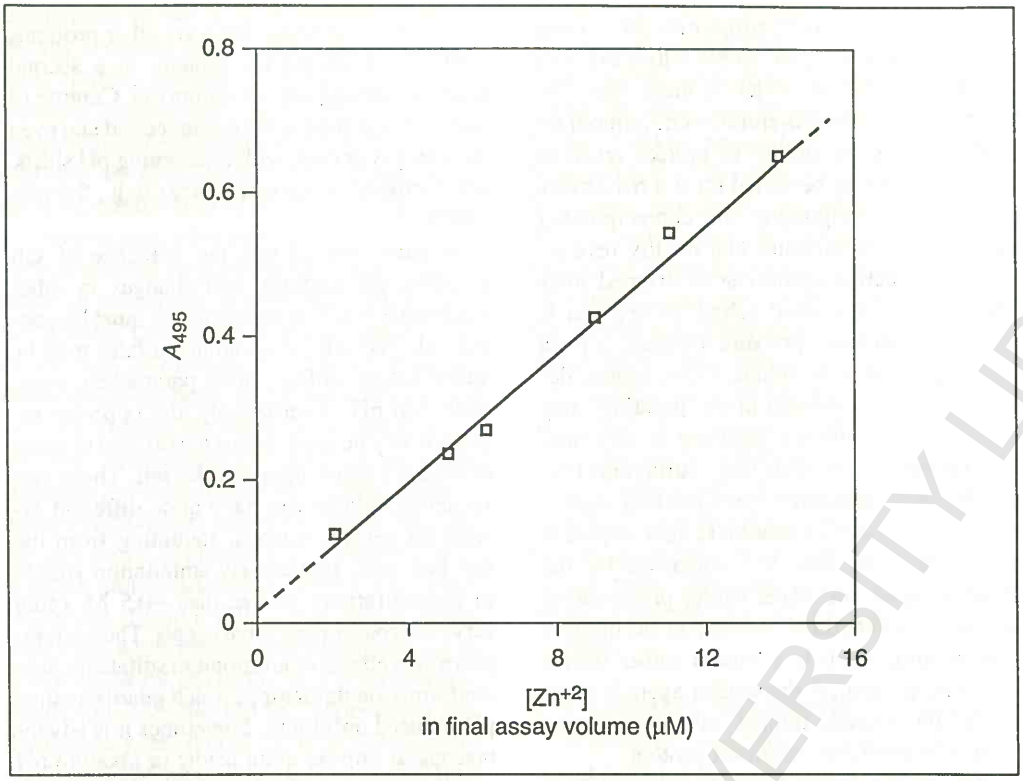


Figure 4.5.12 Calibration plot for zinc ion spectrophotometric analysis using pyridylazoresorcinol reagent.

4. Remove metal ions from precipitate.

   Metal ions may be removed from protein precipitates via cation-exchange resin trapping, or by treating precipitates with a buffer at a pH at which it is expected to dissolve. It is preferable to mix the precipitate with a release buffer first to see if in fact it does dissolve, then test the resin-bead exchanger afterwards. Trapping resins in the Na^+ or K^+ form may be tested; conventional cation resins such as Amberlite IR-120Na^+ usually give adequate results. A specialty chelating resin—Chelex 100 exchange beads (Bio-Rad)—may also be used. The exchange capacities of such resins are 1.8 and 0.4 meq (milliequivalents) per milliliter of wet resin. The resins should not be used in great excess (i.e., >5× resin equivalents per estimated equivalent of metal to be removed) because such resins may also adsorb proteins. Exchange of metal ions out of precipitates may be accelerated by conventional soluble chelating agents (e.g., 10^{-4} M EDTA) or by amino acids such as aspartic acid in the case of Zn^{2+} (Fiabane and Williams, 1977).

5. Perform bioassay for protein of interest.

COMMENTARY

Background Information

In practice, most protein-precipitating agents and methods yield a variety of different products depending upon conditions such as pH, metal ions present in the solution, and/or presence of cosolvents. Many "precipitates" are actually coprecipitates in which molecules of the precipitating agent remain bound to the proteins. How much precipitating agent or ligand becomes bound and thus coprecipitated usually remains unclear until the coprecipitate is analyzed. Precipitation and coprecipitation are most important in "upstream" stages of protein isolation—for crude, dilute, solutions in large volumes up to several liters. However, precipitation is sometimes useful downstream also, as it reduces volume—e.g., excess solvent and buffers may be removed from chromatographic eluates via this technique. Sometimes it is an efficient strategy to precipitate unwanted proteins in the first stages of purification and then capture the sought-for protein in a second precipitation step later.

Some precipitation methods are old, and

BU / OEP / GENERAL UNIVERSITY LIBRARY

some have been developed more recently. Some of the old methods, particularly salting out with ammonium sulfate (which is more than 130 years old), remain as useful as ever. Ammonium sulfate offers the ability to protect sensitive proteins and may be scaled up at a reasonable cost. Some precipitating and coprecipitating agents yield precipitates that readily reverse. Reversibly acting agents can be stripped away later—e.g., ammonium sulfate salting out is reversed by dialysis, pressure filtration, or resin exchange to remove sulfate. Other agents, despite their great powers in precipitating, may be difficult to remove completely, or may catalyze damage to proteins (e.g., sulfhydryl oxidation). Consequently, precipitating agents must be evaluated in relation to later steps that may have to be used to compensate for the disadvantages that some strong precipitating agents carry with them. Sometimes the optimal precipitating agent is a weaker rather than a very strong agent, if the weaker agent is more easily dissociated, trapped out, or dialyzed away to release the sought-for protein.

Besides the agents and individual methods listed in Table 4.5.1, combinations of methods are frequently advantageous. For example, sometimes it is helpful to add modest concentrations of Zn^{2+} to proteins that one is trying to precipitate by another method. Another precipitation technique, TPP (see Basic Protocol 4), is also a "hybrid"—a combination of organic cosolvent precipitation and sulfate kosmotropic precipitation. Precipitation and coprecipitation are often quite protective of proteins, especially in salting-out techniques. However, whether or not protection is achieved frequently depends on the sequence of steps used in each protocol. For example, if salting out or coprecipitation with a particular ligand needs to be carried out at a pH very different from the starting point, a choice must be made between shifting pH first and then adding salt, or alternatively adding salt or ligand first and then readjusting pH. The latter is usually best choice in terms of protecting the protein.

Critical Parameters and Troubleshooting

Salting out

In the simplest version of salting out, desired proteins are precipitated relatively early and the supernatant is set aside or discarded. In this case, unwanted proteins are often left in the supernatant. The supernatant may, however, be readjusted in pH, increased in salt concentration, or both, to bring down yet other proteins, perhaps the sought-for protein, in a second stage or "second cut" of salting out. Capture of sought-for protein activities in second and even third stages or cuts, with intervening pH shifts, are discussed in various sources (e.g., Scopes, 1987).

In each step, or cut, the sequence of salt addition, pH shifting, and changes in other parameters (such as temperature) must be considered. The salt (ammonium sulfate) may be added before shifting other parameters, especially the pH. Alternatively, the opposite sequence may be used: adjust the pH first to some desired set point, then add the salt. These two sequence options can have quite different effects for several reasons, stemming from the fact that salts, particularly ammonium sulfate in concentrations greater than ~0.5 M, often serve as protein-protective agents. The protein-protective effects of ammonium sulfate include conformation tightening, which guards against pH-induced unfolding. Sometimes it is advantageous to impose quite acidic or alkaline pH extremes or a thermal shock, to first precipitate unwanted, denatured proteins, and leave sought-for proteins in the supernatant. However, the amount of salt and when to add it are variables; these need attention in weaving the optimum path for salting out.

Several treatises and textbooks present tables, formulas, and nomograms for relating percent saturation, molar concentration, weight of ammonium sulfate to be added, and dilution factors that are useful for increasing or decreasing amounts of ammonium sulfate so as to arrive at prescribed concentrations of the salt. The numerical relationships are mostly cited for 0° to 30°C, although the solubility of ammonium sulfate in water changes little. Saturated ammonium sulfate in water is close to 4.1 M. Two relationships cited by Scopes (1987) give the number of grams (g) to be added to a liter of solution at 20°C for shifting molarity (M) and percent saturation (S). The first of these equations is for shifting to a molarity M_2 starting at molarity M_1:

$$g = \frac{553(M_2 - M_1)}{4.05 - 0.3M_2}$$

The second equation is for shifting to a percent saturation S_2 starting at percent saturation S_1.

$$g = \frac{553(S_2 - S_1)}{100 - 0.3S_2}$$

These relationships pertain to solutions of

PCU IASRI LIBRARY

ammonium sulfate in water. Proteins that are already in diverse buffers and that contain other salts, in unknown concentrations, throw off the accuracy of nomograms and formulas for shifting saturation percentages and molarities. Nevertheless, these are useful for estimating added salt concentrations to be used and for reporting.

Neutral salts other than ammonium sulfate, such as sodium or potassium chloride, may be tested in attempting to precipitate the sought-for protein (and later in attempting to split excess salt away—i.e., reverse the process). However, ammonium sulfate is usually the best primary choice, not because of tradition but because the sulfate anion, SO_4^{-2}, is one of the strongest yet most benign Hofmeister kosmotropes (*UNIT 8.4*) over nearly the whole "biological" pH range, approximately from pH 2 to 10 (Collins and Washabaugh, 1985). The orthophosphate dianion, HPO_4^{-2} is also a strong, inexpensive kosmotrope and salting-out agent. However the pK of $H_2PO_4^-$ is 7.2; accordingly, below pH ~6.5 the $H_2PO_4^-$ monoanion predominates even though it is a considerably weaker kosmotrope. Other anions such as fluoride, F^-, are good kosmotropes but are very toxic. For a combination of desirable properties, including least expense, the sulfate systems remain the most convenient, reliable, and most tested. One disadvantage of sulfate arises in the case of proteins requiring calcium ion, Ca^{2+}. Because calcium sulfate and calcium phosphate are extremely insoluble, both SO_4^{-2} and HPO_4^{-2} precipitate Ca^{2+}. Hence, for calcium-requiring proteins, salting out may have to be developed with salts other than sulfates—e.g., acetates are a possibility.

Isoionic precipitation

Amberlite resins IR-120 (a cation exchanger) and IRA-400 (an anion exchanger), in the form of 6% or 8% cross-linked beads, are the basis of the original Dintzis mixed-bed exchangers. The equivalent Dowex exchangers, Dowex 1 and Dowex 50 series, in the form of 6% to 8% cross-linked resins, serve equally well but are not superior.

Anion and cation exchangers have differing exchange capacities per settled milliliter of wet resin. IR-120H^+ has ~1.8 meq/ml of capacity, and IRA-400OH^- has ~1.2 meq/ml. Accordingly each 100 ml of mixed exchanger may be made by mixing 40 ml of the cation exchanger in the H^+ form with 60 ml of anion exchanger in the OH^- form, a volume ratio inversely proportional to their intrinsic exchange capacity ratio. Mixed-bed exchangers may be stored in the cold (do not freeze them) for long periods without self-neutralization.

Mixed-bed exchangers, such as the MB-1 Amberlite series, may be purchased premixed. However it is recommended that the exchanger be mixed in the laboratory to ensure that resins have been "exercised" (see Basic Protocol 2, steps 2 and 3) and leached of any polymer hydrolysis products. Anion exchangers based on various quaternary amines are especially susceptible to self-hydrolysis (detectable by an amine odor) upon storage in the OH^- form. Commercial ready-mixed bed resins usually incorporate dye chromogens (acid-base indicator molecules), which show (by a blue to amber color change) when their capacity has been titrated. Such resins may be usable with some protein molecules, but if there is any leaching out of resin materials, the eluants will be strongly colored, indicating contamination of proteins.

Two-carbon organic cosolvent precipitation

In Basic Protocol 3, no general criterion can be put forth for choosing between acetone and ethanol, for deciding how pH should be adjusted versus the isoionic point of the protein, or for determining the exact salt concentration to be used. There exists a large body of literature describing empirical, experimental means for C_2 cosolvent–based isolation of dozens of proteins. A few guidelines may be useful. Ethanol was the C_2 cosolvent used for the isolation of most blood-plasma proteins. Many respiratory enzymes have been isolated with a step involving an acetone cosolvent. In systems where lipids and triglycerides are troublesome, many workers have used acetone cosolvents, sometimes at nearly 100% acetone, because acetone and acetone-water mixtures extract lipid materials. Acetone in such applications functions as an extractant as well as a protein precipitating agent. Proteins and whole-tissue agglomerates that have been rendered nearly lipid-free by acetone, then lyophilized, are referred to as "acetone powders."

A review with many references was recently written by Rothstein (1994), which includes discussion of the seminal papers by Cohn and coworkers. With a number of proteins, a complex relationship exists between salts, the C_2 organic cosolvents, the proteins, (which bind salts to some extent), and the pH. These parameters operate to make organic cosolvent precipitation quite empirical.

Methanol and the two-carbon organic cosolvents appear to be able to penetrate protein

PRINCETON UNIVERSITY LIBRARY

molecules and compete with water for sites inside. However C_4 and larger cosolvents are largely excluded from the inside of proteins if the proteins are kept compact and folded. Accordingly, it is unlikely that the larger cosolvent molecules, such as C_4 alcohols, behave congruently with the C_1 and C_2 cosolvents for most proteins.

C_4 and C_5 organic cosolvent precipitation

Some C_4 organic cosolvents, particularly *t*-butanol and *n*-butanol, are able precipitating and extracting agents for proteins. They are useful in "upstream" stages of isolation, starting with crude solutions. They are variously used in combination with osmolytes and salting-out agents such as ammonium sulfate. Conditions are adjusted in such a way that separated layers of cosolvent are obtained, along with an aqueous layer and a layer of precipitated protein. The separated liquid layers act as extraction and partitioning systems as well as protein precipitation systems. Precipitated proteins usually appear at the interface between liquid layers. Because some of the C_4 and C_5 cosolvents act either as kosmotropes or in parallel to kosmotropes like the sulfate anion, these cosolvents frequently are protective of aqueous proteins even at normal temperatures, 20° to 30°C. They induce protein-molecule conformation tightening, from which the protein-molecule protection partly originates. The similarity between sulfate interaction with proteins and C_4/C_5 cosolvent interaction with proteins does not, however, have the same origin. Sulfate exerts electrostatic forces toward many proteins (Chakrabarti, 1993), whereas *t*-butanol exerts much less electrostatic-based interaction, perhaps none at all. A rough model of the effect of the *t*-butanol cosolvent is that it behaves like a neutral osmolyte in water.

It is not always strictly necessary to choose *t*-butanol as the cosolvent. However, *t*-butanol has some advantages. First, it is relatively inexpensive. Only ~0.3 to 0.5 ml of *t*-butanol is required per milliliter of starting aqueous protein solution, after the optimal amount of salt has been added. When *t*-butanol is used, smaller amounts of ammonium sulfate are usually required for a given precipitation, relative to the amounts needed in conventional salting out. Second, *t*-butanol is much less volatile than C_2 cosolvents, especially in water-dominated systems. Hence it poses less of a fire hazard and less hazard from inhalation. Third, *t*-butanol is, in general, far less denaturing at 20° to 30°C temperatures than C_2 cosolvents. Fourth, *t*-butanol behaves similarly to kosmotropes and osmolytes, protecting proteins. Finally, the extractive and partitioning capacities of three-phase partitioning (TPP) using C_4 and C_5 organic cosolvents (see Fig. 4.5.6) tend to remove pigments, lipids, and amino acids from proteins.

C_4 and C_5 organic cosolvents in the aqueous phase produced by techniques like TPP exert themselves as osmolytes and crowding agents, up to a point, when they reach concentrations of several percent. However, the main differences between the C_4/C_5 organic compounds (of the types discussed here—e.g., butanols and pentanols) and water-soluble osmolytes such as sugars are: (1) osmolytes like sugars and amino acids tend not to form a separate organic phase that goes on to act as an extraction phase for lipids, pigments, and other nonprotein material, whereas the butanols and pentanols do; and (2) hydrocarbonaceous C_4 and C_5 compounds bind to some extent to proteins that have a high degree of nonpolar, hydrophobic character. This latter property causes such proteins to position themselves as illustrated in Figure 4.5.6, floating above the aqueous layer. When the protein precipitate (i.e., coprecipitate) is redissolved in buffer with low salt, then shifted away from the precipitative pH (usually upward ~2 to 3 pH units), in most instances the C_4/C_5 compounds dissociate. In contrast, conventional osmolytes, especially sugar, tend not to bind to proteins.

In some versions of TPP, the system is operated in two stages. In the first stage, the precipitate containing unwanted proteins that forms at the initial pH is discarded. The aqueous phase is then shifted in pH, and/or more salt is added, giving a second precipitate containing the sought-for protein. In the example described in Lovrien et al. (1987), the enzyme α-amylase is isolated in this way. Many sorts of variations can be made in pH, salt concentration, and other parameters.

Protein-exclusion and crowding agents

Exclusion agents and osmolytes have high affinities for water and compete with proteins in solution for water. The crowding action, aided by any degree of affinity of protein molecules for one another (protein-to-protein association) promotes protein precipitation. Thermodynamic pushing occurs because hostile environments imposed on proteins in osmolytes and and in crowded solution raises protein chemical potential in a positive direction. This is then relieved by protein precipitation, which

PCU IAS LIBRARY

enables protein molecules to exist with lowered chemical potential.

Crowding and exclusion agents frequently act upon conformationally motile (loose), water-penetrated proteins before overt precipitation. Conformationally loose protein molecules are also "squeezed" on by these agents, promoting protein molecule tightening and sometimes promoting an ordered protein conformation. Proteins thus "squeezed on" often become protected, sometimes dramatically. Moreover if their conformation is forced into good order, protein molecules may crystallize instead of simply being precipitated in amorphous form. Thus, the osmolyte MPD (2-methyl-2,4-pentanediol) is often used with success for protein crystallization. MPD crowds, excludes, and squeezes to push proteins into a narrow or even single conformation necessary for arriving at the crystalline state.

As a general rule, osmolytes and crowding agents do not bind strongly to proteins. Hence, ideally, precipitated proteins are free of such agents. However in fact, some agents bind rather weakly, and appear in the precipitate (which is in fact a coprecipitate). Accordingly, selection of the agent may depend on the ease or difficulty of removing trapped agent from protein precipitates if in fact the desired redissolved protein needs be stripped of precipitating agent.

Neutral polymers such as polyethylene glycol (PEG) and dextran, which do not affect electric charge, might be expected to act toward proteins independently of electrostatic parameters such as protein net charge and salt concentration. However, the opposite often occurs: i.e., buffer, salt, and pH-dependent protein charge frequently sharpen and either greatly enhance or sometimes depress protein precipitation via PEG. Salts in concentrations quite below normal kosmotropic concentrations—i.e., 0.01 to 0.5 M salts—have a large influence, decreasing chymotrypsin solubility and increasing its precipitation by PEG (Miekka and Ingham, 1978). In other examples, chaotropic anions and "Hofmeister-neutral" anions (e.g., chloride) increase protein solubility and depress precipitation. Readers may profit by consulting some of the papers referred to here, particularly to Miekka and Ingham (1978) to review how large and variable these "electrostatic" effects may be.

Neutral polymers, acting as steric-exclusion precipitants, and the monomers of such polymers, acting as osmolytes, are far from equivalent in protein-precipitating power when used in equal concentrations. The polymers appear to be considerably most effective. For example, Laurent (1963) investigated dextran polymer and its monomer glucose as precipitants for albumin and fibrinogen. Dextran polymer considerably outstripped glucose in precipitation power when each reagent was used at a concentration of ~0% to 10% (w/v). On the other hand, concentrations of proteins for achieving protein precipitation using neutral polymers usually must be from 1% to 5% (w/v). Hence the neutral polymers are considerably less sensitive and less effective at low protein concentrations than some of the other methods discussed in this unit—e.g., the matrix-stacked ligand technique (see Basic Protocol 9 and discussion below).

Neutral polymers are equivalent to neutral polymers used in aqueous two-phase partitioning (Walter and Johansson, 1994) and as "precipitants" in protein crystallization (McPherson, 1982). PEG is the dominant polymer used in two-phase partitioning and as a precipitant for protein crystallization. PEG polymers in several molecular-weight ranges—low, medium, and high—are available from Sigma and other suppliers. Choice of PEG molecular weights must be balanced between two considerations. Higher-molecular-weight polymers tend to be more effective precipitants for some proteins (i.e., able to precipitate albumins out of solution in ~5% to 15% PEG concentration ranges; Atha and Ingham, 1981). Lower-molecular-weight PEG polymers of molecular weight <3000 to 4000 generally must be used in 20% to 50% concentrations to effectively precipitate a number of proteins that have been studied. However, the practical problem of ridding precipitated proteins of trapped PEG polymers is eased with low-molecular-weight PEGs because they dialyze away relatively easily. Proteins generally are more sensitive to precipitation by high-molecular-weight PEGs than low molecular-weight-PEGs, frequently precipitating, if indeed they do so, in 3% to 20% concentrations of large PEGs (molecular weight greater than ~4000; Atha and Ingham, 1981). The downside of high-molecular-weight PEGs is that, if trapped, these PEGs become difficult to separate from coprecipitated proteins by simple dialysis. Also, high-molecular-weight PEG solutions tend to be quite viscous, whereas oligomeric low-molecular-weight PEGs are easier to handle and mix.

Alkane ether polymers (trade name Ucon; Union Carbide; see *SUPPLIERS APPENDIX*), which are copolymers of ethylene and propylene ox-

POLYCENTRAL LIBRARY

ide related to PEGs, go into and out of solution with marked temperature dependency (i.e., they "phase-separate"; Albertsson and Tjerneld, 1994). Accordingly, with the use of such polymers, manipulation of temperatures in initial precipitation and in redissolving protein precipitates may afford means for balancing these two problems.

Synthetic and semisynthetic polyelectrolytes

Polyelectrolytes have been synthesized as copolymers with nonionic groups to help alleviate problems that arise with homopolymers such as polyacrylic and polymethacrylic acid. The latter homopolyelectrolytes, when ionized, are so strongly coprecipitating that they are often poorly selective. Hence, copolymers of acrylic acid and methacrylic acid with a nonionic methylmethacrylate ester side function as agents with softened coprecipitating power but increased selectivity. A series of copolymers of this type, with varying ratios of ionic and nonionic groups, are named "Eudragit" agents, and are available commercially from Rohm GmbH (see *SUPPLIERS APPENDIX* and Mattiasson and Kaul, 1994). Copolymerization of two and even three separate vinyl monomers provide polyampholytes of myriad kinds, differing in ratios of vinyl monomers. Synthetic polyampholytes take on net positive, net negative, or zero charge, depending on the pH (like proteins, which are also polyampholytes). The advantage of such compounds is that after coprecipitating proteins, the synthetic polyampholyte can be split away from the protein by adjusting the pH of the mixture to the isoionic pH of the polyampholyte, which will precipitate it separately (Patrickios et al., 1994). Many of the polymers that have been discussed so far—e.g., carboxymethylcellulose and polymethacrylic acid anion—are negatively charged above the pK of their side-chain carboxyl groups, ~pH 5. Such polymers should primarily be used for precipitating proteins that are cationic—i.e., in pH ranges below the target protein's isoelectric point. However, it may be necessary to experiment with synthetic polyelectrolytes which are cations and which are therefore primarily aimed at precipitating anionic proteins in pH ranges impractical for synthetic polyanions. The most common, readily available neutral-cationic synthetic polyelectrolyte of this type is polyethyleneimine (PEI). The H^+-titratable nitrogen groups of PEI are integral parts of its backbone, not relegated to side chains. PEI is very basic and exists as a polycation over a very wide range of pH. PEI is also a powerful coprecipitating agent for proteins. Table 4.5.2 lists a small number of polyelectrolytes, several of which are sold by Aldrich and Sigma. References in Table 4.5.2 should enable the reader to survey conditions used for coprecipitation of various enzymes.

Metallic and polyphenolic heteropolyanions

There is considerable variability but lack of much quantitative information concerning recovery and bioactivity measurements of most proteins that have been precipitated with heteropolyanions, in cases where their activities must be recovered. The great acidity and also the oxidizing power of some of these agents—e.g., $HClO_4$—tends to seriously denature many proteins. Deamidation of proteins sometimes follows strong acidification, as well as unwanted hydrolysis and conformational unfolding. However, weaker acids such as tannic acid (a mixture of several phenolic products) are less denaturing, and yield intact active enzymes and other proteins after tannate precipitation at neutral pH. Accordingly, choosing among heteropolyanion precipitation agents depends on the use to which the protein precipitates are to be put. For very rapid, complete precipitation of unwanted proteins down to ~50 µg/ml (Bensadoun and Weinstein, 1976), in cases where destruction of proteins is acceptable, $HClO_4$ in 1% to 5% concentration is used. Picric acid is almost equally effective, but it is highly colored and hence interferes in analytic procedures requiring spectrophotometric measurement. As with polyelectrolyte precipitation, which agent to use may depend on the ease versus difficulty of removing such agents from proteins after coprecipitation.

Examples of precipitation of diverse proteins by heteropoly anions at mildly acidic pH (i.e., pH 4 to 6) are described in two papers. Sternberg (1970) coprecipitated ovalbumin, hemoglobin, pepsin, and amyloglucosidase from protein crudes using four heteropoly anions. Sternberg removed the inorganic anions by dialysis, ultrafiltration, and ion exchange, recovering fair to good activities in some cases. Astrup et al. (1954) worked with six heteropolyanions in staged purification of serum proteins including α- and γ-globulins.

Tannic acids are polyphenolic compounds (sometimes conjugated with one or more saccharides) from plant extracts. They are quite diverse in structure, depending on the plant source. Sigma sells a crude tannic acid that is convenient to start with. An advantage to using tannates as precipitants is that they function in

RCU IVSTITVT LIBRARY

neutral pH regions, pH ~4 to 8. Hence, severe acid and hydrolytic conditions, often necessary with the metallic heteropolyanions, may be avoided with tannates, which are all-organic anions. Little detail is known concerning how tannate precipitation works. Tannates are sometimes effective in precipitating proteins on a broad front, and inclusively, but in some systems fail to precipitate proteins to which they are bound (Hagerman and Butler, 1991). Amounts of tannic acid and of tannate needed for precipitation of diverse proteins range from approximately half as much by weight as the sought-for protein, down to quite smaller amounts of tannin; these quantities are also quite dependent on pH. Hagerman and Butler (1978, 1991) survey principal parameters for controlling precipitation of individual proteins such as albumin and lysozyme. They recommend quebracho tannin and commercial tannic acids as initial agents for attempting precipitation of mixtures of proteins. Although there are gaps in the precipitating power of tannate compounds for proteins, they provide alternatives for scaleable, reversible procedures under conditions where the metallic heteropolyanions (strong acid) are too severe. The strong spectrophotometric absorption spectra of tannins when reacted with Folin phenol reagent or with ferric chloride, both of which generate characteristic colors for phenols, afford convenient means for quantitating tannates in coprecipitates and in soluble remnants (Hagerman and Butler, 1978).

Hydrophobic ion-pairing entanglement ions

It may seem surprising that ligands such as those used in Basic Protocol 8—especially detergents—are efficient precipitating agents; these reagents are normally used in solubilizing and denaturing proteins. Detergent anions (i.e., dodecyl sulfate) acting as protein coprecipitants were first studied at length by Putnam and Neurath (1944) for seven proteins. Since then, the detergent/protein and subsequent complex/complex association reactions have sometimes been called hydrophobic ion-pairing reactions. A rather strict detergent-anion to protein-cation association occurs in 1:1 stoichiometry to form ion pairs between charged groups; simultaneously an alkane/alkane association occurs between the hydrophobic tails. The two sets of forces strongly pull the complexes together and out of the solution to agglomerate and form the visible coprecipitate. Protein-molecule to protein-molecule association may also occur, but there are rather few detailed studies of this. The ion/ion association, which is the electrostatic component of HIP, is dependent on and adjusted by the pH, because the pH and the number of protein H^+–titratable side chains determine the number of net cationic charges on the global protein molecule—i.e., Z_{H^+}. This, in turn, determines the number of ligand anions expected to bind when coprecipitates are formed. Because strong anions from detergents usually bind strongly, there is seldom a need to add extra ligand (detergent) to provide for large amounts of free (unbound) ligand. Between 60% and 95% of the total detergent anions added under the aforementioned conditions are likely to be bound for 10-carbon, 12-carbon, and larger detergent sulfates. However, short-chain detergent anions—e.g., hexyl and octyl sulfates—are usually poor HIP precipitants.

Matrix-stacking ligand coprecipitation

In organic anion–matrix ligands, the rigid azoaromatic sulfonates are able to stack their tail groups to force association and coprecipitation. These reagents have several similarities to flexible detergent coprecipitating anions (HIP agents; see discussion above). As in the case of the flexible detergents, matrix-ligand coprecipitation is sharply dependent on protein electrostatic charge, thence on pH, on protein side-chain composition, and on the isoionic point. In contrast to the detergent anions, whose coprecipitating powers are mainly dependent on alkane-chain length, matrix-stacking ligands are dependent on organic structural detail and placement of groups, not simply on size. The placement of short alkane "reinforcement" substituents and the location of sulfonate anions on aromatic rings, as well as the location of tautomerizing hydroxyl groups, make stacking-ligand structural character more complicated and unpredictable than that of flexible-detergent ligands. In practice, one may need to experiment with a variety of matrix ligands to find an optimal kind for selectively coprecipitating each protein.

It appears that matrix-stacking ligands need their hydrophobic tail groups to extend some distance away (greater than ~10 to 15 Å) from the sulfonated head group (Matulis and Lovrien, 1996). This requirement presumably maximizes tail-tail stacking and hydrophobic association, distancing them so as not to interfere with sulfonate group–protein cation association. In any case, functions of matrix ligands are sensitive to placement of substituents such as alkane groups on the azoaromatic rings, as

POLYTECHNIC UNIVERSITY LIBRARY

well as the number of such substituents. In practice, one expects to initially experiment with ~3 to 5 such ligands to find a ligand suitable for investigation in more detail. Because matrix ligands generally are quite protective of proteins with which they coprecipitate, it may not be necessary to always develop the coprecipitation stage of protein isolation at cold temperatures. Indeed, the method often interfaces well with a heat-shock step to clear out unwanted proteins susceptible to heat shocking. Pilot experiments therefore can generally be carried out at ordinary room temperatures.

Matrix ligands strongly coprecipitate proteins. However, overall isolation processes for most proteins usually need an exit. The protein and ligand should be released from one another by resin-exchange trapping as described in Basic Protocol 2. Hence, it may occur that somewhat more weakly coprecipitating, but more easily released ligands may be optimal in the procedures, so as to allow proteins to dissolve, ready for assay or the next isolation steps.

Di- and trivalent metal cation precipitation

Metal cations exert themselves in three ways to precipitate proteins. First, their charges add to protein molecule net charge, enhancing binding of coprecipitating ligand ions such as matrix or detergent sulfate anions. Second, metal ions draw in protein side chains and lysine amines, as well as carboxylate and histidyl side chains, to fill metal ion orbitals with their electron pairs. Protein molecules thus act as large chelating agents in a manner analogous to the behavior of EDTA and other multidentate chelators and become tightened and sometimes protected because of the chelation to M^{2+}. Finally, M^{2+} and M^{3+} ions can become complexed with buffer ions and other electrolytes chosen by the researcher. When that happens, actual net charge and remaining metal ion ligand positions available to the protein may be adjusted to modulate protein precipitation. Metal ion complexes of M^{2+} and M^{3+} may actually be made neutral, or even anionic, depending on presence of chloride, acetate, or citrate, and their concentrations (in the 10^{-3} to 10^{-1} M range); this phenomenon is also dependent on pH. Cotton and Wilkinson (1988) cite $Zn(H_2O)_6^{+2}$, $ZnCl(H_2O)_5^{+1}$, $ZnCl_2(H_2O)_4$, $ZnCl_4^{-2}$, and $ZnCl_4(H_2O)_4^{-2}$ as examples. Cobalt and a number of other transition metal cations complex with ammonia and water molecules to make a series of coordinate cations of considerably larger size than the parent Co^{2+} but which still carry the central metal ion's charge. On the other hand, addition of CN^- or SCN^- generates a series of complexes such as $Fe(CN)_6^{-3}$ which are strong, very stable anions. A number of metal cations—e.g., Ni^{+2} and Zn^{+2}—are capable of switching configuration from square planar to tetrahedral to octahedral, depending on coordinate ligands, chelating agent, and solvent conditions. Some metal cations useful in protein precipitation are amphoteric, switching water molecules and OH^- ligands into and out of metal orbital positions depending on pH. Metal cations such as Al^{3+} strongly bind oxo-ligands, including water molecules and hydroxyl (OH^-) to form metallic gels that settle or centrifuge down and capture proteins by adsorption (Collingwood et al., 1988). This mode of precipitation is often called flocculation. Flocculated products tend to have uncertain composition, but have the advantage of being dense and separable by low-speed centrifugation or simply by settling.

If the choice of metal ion is otherwise not obvious, the first choice as protein coprecipitant is the zinc ion. Zn^{2+} is of intermediate strength in chelation-coordination reactions with respect to the kinds of ligands offered by amine, carboxylate, and imidazole side chains. Therefore it can be "reversed"—i.e., removed from coprecipitates. Zn^{2+} rather rapidly exchanges ligands that coordinate with it (Berg and Shi, 1996). More importantly, zinc ions do not promote electron transfer and oxidation-reduction reactions, whereas many of the transition metals—e.g., iron, copper, and nickel ions—indeed do so. Hence Zn^{2+} is fairly safe to use for precipitating proteins with sensitive sulfhydryl (–SH) groups. Other metal ions–e.g., Cu^{2+} and Ni^{2+} promote oxidation, usually irreversibly, and therefore can be rather destructive. Zn^{2+} in the form, for example, of zinc chloride, is also the probable first choice to use as an "assist" agent in conjunction with other precipitants. Atha and Ingham (1981) show excellent examples of the ability of $ZnCl_2$ to decisively decrease the concentration of PEG crowding agent necessary to precipitate out human serum albumin from a 0.5% solution. Zn^{2+} appears to tighten the albumin molecule's conformation, easing PEG's task in precipitating the protein.

Literature Cited

Albertsson, P.A. and Tjerneld, F. 1994. Phase diagrams. *Methods Enzymol.* 228:3-13.

Alner, D.J., Creczek, J.J., and Smeeth, A.G. 1967. Precise pH standards from 0° to 60°. *J. Chem. Soc. A* 1205-1211.

RCU IAS CENTRAL UNIVERSITY LIBRARY

Astrup, T., Birch-Andersen, A., and Schilling, K. 1954. Fractional precipitation of serum proteins by means of specific anions. *Acta Chem. Scand.* 8:901-908.

Atha, D.H. and Ingham, K.C. 1981. Mechanism of precipitation of proteins by polyethylene glycols. *J. Biol. Chem.* 256:12108-12117.

Berg, J.M. and Shi, Y. 1996. The galvanization of biology: A growing appreciation for the roles of zinc. *Science* 271:1081-1085.

Bensadoun, A. and Weinstein, D. 1976. Assay of proteins in the presence of interfering materials. *Anal. Biochem.* 70:241-250.

Berdick, M. and Morawetz, H. 1954. The interaction of catalase with synthetic polyelectrolytes. *J. Biol. Chem.* 206:959-971.

Bickle, T.A., Pirotta, V., and Imber, R. 1977. A single, general procedure for purifying restriction endonucleases. *Nucl. Acids Res.* 4:2561-2572.

Budavari, S. (ed.) 1989. The Merck Index, 11th ed. Merck & Co., Rahway, N.J.

Chakrabarti, P. 1993. Anion binding sites in protein structures. *J. Mol. Biol.* 234:463-482.

Cheng, K.L., Ueno, K., and Imamura, T. (eds.) 1992. CRC Handbook of Organic Analytical Reagents, pp. 219-226. CRC Press, Boca Raton, Fla.

Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.J., Ashworth, J.N., Melin, M., and Taylor, H.L. 1946. Preparation and properties of serum and plasma proteins. IV. A system for separation into fractions of the protein and lipoprotein components of biological tissue and fluids. *J. Am. Chem. Soc.* 68:459-475.

Cohn, E.J., Gurd, F.R., Surgenor, D.M., Barnes, B.A., Brown, R.K., Derouaux, G., Gillespie, J.M., Kahnt, F.W., Lever, W.F., Liu, C.H., Mittelman, D., Mouton, R.F., Schmid, K., and Uroma, E. 1950. A system for separation of the components of human blood: Quantitative procedures for separation of the protein components of human plasma. *J. Am. Chem. Soc.* 72:465-474.

Collingwood, T.N., Daniel, R.M., and Langdon, A.G. 1988. An M(III) facilitated flocculation technique for enzyme recovery and concentration. *J. Biochem. Biophys. Methods* 17:303-310.

Collins, K.D. and Washabaugh, M.W. 1985. The Hofmeister effect and the behavior of water at interfaces. *Q. Rev. Biophys.* 18:323-422.

Conroy, M.J. and Lovrien, R.E. 1992. Matrix coprecipitating and cocrystallizing ligands (MCC ligands) for bioseparations. *J. Crystal Growth* 122:213-222.

Cotton, F.A. and Wilkinson, G. 1988. Advanced Inorganic Chemistry, 5th ed., pp. 606-608. Wiley-Interscience, New York.

Dubin, P., Ross, T.D., Sharma, I., and Yegerlehner, B. 1987. Coacervation polyelectrolyte-protein complexes. *In* ACS Symposia Series 342 (W. Hinze and D. Armstrong, eds.), pp. 162-169. American Chemical Society, Washington, D.C.

Edsall, J.T. and Wyman, J. 1958. Biophysical Chemistry, pp. 591-662. Academic Press, New York.

Felix, K. 1960. Proamines. *Adv. Protein Chem.* 15:1-20.

Fiabane, A.M. and Williams, D.R. 1977. The Principles of Bio-Inorganic Chemistry. *In* Chemical Society Publication no. 31, pp. 32-48. The Chemical Society, London.

Fink, A.L. 1995. Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 24:495-522.

Gurd, F.R.N. and Wilcox, P.E. 1956. Complex formation between metallic cations and proteins, peptides and amino acids. *Adv. Protein Chem.* 11:311-427.

Hagerman, A.E. and Butler, L.G. 1978. Protein precipitation method for the quantitative determination of tannins. *Agric. Food Chem.* 26:809-812.

Hagerman, A.E. and Butler, L.G. 1991. Tannins and lignins. *In* Herbivores: Their Interactions with Secondary Plant Metabolites, Vol. 1 (G. Rosenthal, ed.) pp. 355-388. Academic Press, New York.

Herzfeld, J. 1996. Entropically driven order in crowded solutions: From liquid crystals to cell biology. *Acc. Chem. Res.* 29:31-37.

Hidalgo, J. and Hansen, P.M.T. 1971. Selective precipitation of whey proteins with carboxymethylcellulose. *J. Dairy Sci.* 54:1270-1274.

Hogness, T.R. and Johnson, W.C. 1954. Qualitative Analysis and Chemical Equilibrium, 4th ed., pp. 567-598. Henry Holt Co., New York.

Hunt, J.B., Neece, S.H., and Ginsburg, A. 1985. The use of 4-(2-pyridylazo)resorcinol in studies of zinc release from *E. coli* aspartate transcarbamylase. *Anal. Biochem.* 146:150-157.

Jendrisak, J. 1987. Use of polyethyleneimine in protein purification. *In* Protein Purification: Micro to Macro (R. Burgess, ed.) pp. 75-97. Alan R. Liss, New York.

Jendrisak, J.J. and Burgess, R.R. 1975. A new method for the large scale purification of wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 14:4639-4645.

Kressman, T.R. 1957. Commercial materials. *In* Ion Exchangers in Organic and Biochemistry (C. Calmon and T. Kressman, eds.) pp. 107-109. Interscience Publishers, New York.

Laurent, T.C. 1963. The interaction between polysaccharides and other macromolecules. *Biochem. J.* 89: 253-257.

Lovrien, R., Goldensoph, C., Anderson, P.C., and Odegaard, B. 1987. Three phase partitioning (TPP) via *t*-butanol: Enzyme separations from crudes. *In* Protein Purification: Micro to Macro (R. Burgess, ed.), pp. 131-148. Alan R. Liss, New York.

Mattiasson, B. and Kaul, R. 1994. One-pot purification by process integration. *Biotechnology* 12:1088-1089.

BIOMEDICAL CENTRAL LIBRARY

Matulis, D. and Lovrien, R. 1996. Coprecipitation of proteins with matrix ligands: Scaleable protein isolation. *J. Mol. Recognit.* In press.

McPherson, A. 1982. Preparation and Analysis of Protein Crystals, pp. 102-108. John Wiley & Sons, New York.

Miekka, S.I. and Ingham, K.C. 1978. Influence of self-association proteins on their precipitation by polyethylene glycol. *Arch. Biochem. Biophys.* 191:525-536.

Morawetz, H. and Hughes, W.L., Jr. 1952. The interaction of proteins with synthetic polyelectrolytes. I. Complexing of serum albumin. *J. Phys. Chem.* 56:64-69.

Morton, R.K. 1950. Separation and purification of enzymes associated with insoluble particles. *Nature* 166:1092-1095.

Niederauer, M.Q., Suominen, I., Rougvie, M.A., Ford, C.F., and Glatz, C.E. 1994. Characterization and polyelectrolyte precipitation of β-galactosidase containing genetic fusions of charged polypeptides. *Biotechnol. Prog.* 10:237-245.

Patrickios, C.S., Hertler, W.R., and Hatton, T.A. 1994. Protein complexation with acrylic polyampholytes. *Biotechnol. Bioeng.* 44:1031-1039.

Perrin, D.D. and Dempsey, B. 1974. Metal-ion buffers. *In* Buffers for pH and Metal Ion Control, pp. 94-102. Chapman & Hall, London.

Putnam, F.W. and Neurath, H. 1944. The precipitation of proteins by synthetic detergents. *J. Am. Chem. Soc.* 66:692-697.

Rothstein, F. 1994. Differential precipitation of proteins. *In* Protein Purification Process Engineering (R.G. Harrison, ed.) pp. 115-208. Marcel Dekker, New York.

Sandell, E.B. and Onishi, H. (eds.) 1978. Photometric Determination of Traces of Metals, 4th ed. John Wiley & Sons, New York.

Scopes, R.K. 1987. Protein Purification: Principles and Practice, 2nd ed., pp. 45-54. Springer-Verlag, New York.

Sela, M. and Katchalski, E. 1959. Biological properties of poly-α-amino acids. *Adv. Protein Chem.* 14:391-478.

Shibata, T., Cunningham, R.P., and Radding, C.M. 1981. Homologous pairing in genetic recombination. *J. Biol. Chem.* 256:7557-7564.

Singh, R.K. 1995. Precipitation of biomolecules. *In* Bioseparation Processes in Foods (R.K. Singh and S.S. Rizvi, eds.) pp. 139-173. Marcel Dekker, New York.

Sternberg, M.Z. 1970. The separation of proteins with heteropolyacids. *Biotechnol. Bioeng.* 12:1-17.

Sternberg, M. 1976. Purification of industrial enzymes with polyacrylic acids. *Process Biochem.* 11:11-12.

Sternberg, M. and Hershberger, D. 1974. Separation of proteins with polyacrylic acids. *Biochim. Biophys. Acta* 342:195-206.

Tanford, C. 1961. Multiple equilibria. *In* Physical Chemistry of Macromolecules, pp. 561-564. John Wiley & Sons, New York.

Timasheff, S.N. 1992. Water as ligand: Preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* 31:9857-9864.

Walter, H., and Johansson, G. 1994. Aqueous two-phase systems. *Methods Enzymol.* 228: 1-17.

Whitaker, D.R. 1953. Purification of *M. verrcaria* cellulase. *Arch. Biochem. Biophys.* 43: 253-267.

Contributed by Rex E. Lovrien and Daumantas Matulis
University of Minnesota
St. Paul, Minnesota

RCU IASI / CENTRAL UNIVERSITY LIBRARY

PRODUCTION OF RECOMBINANT PROTEINS

5

INTRODUCTION 5.0.1

5.1 Production of Recombinant Proteins in *Escherichia coli* **5.1.1**
General Strategies for Gene Expression in *E. coli* 5.1.1
Specific Expression Strategies 5.1.2
Summary 5.1.6
Literature Cited 5.1.6

5.2 Selection of *Escherichia coli* Expression Systems **5.2.1**
Basic Protocol 1: Expression Under Control of p_L Promoter: Temperature Induction 5.2.2
Basic Protocol 2: Expression Under Control of *trp* Promoter: Chemical Induction 5.2.3
Alternate Protocol 1: Expression Under Control of *lac*/tac Promoters 5.2.4
Alternate Protocol 2: T7 RNA Polymerase/Promoter Expression System 5.2.5
Support Protocol 1: Transformation by Electroporation 5.2.5
Support Protocol 2: Plasmid Stability Analysis 5.2.7
Support Protocol 3: Strain Storage 5.2.8
Support Protocol 4: Preparation of Samples for Analysis by SDS-PAGE 5.2.9
Support Protocol 5: Solubility Analysis 5.2.10
Support Protocol 6: Preparation of Periplasmic Extracts 5.2.11
Support Protocol 7: Preparation of Extracellular Medium Samples 5.2.12
Reagents and Solutions 5.2.12
Commentary 5.2.14

5.3 Fermentation and Growth of *Escherichia coli* for Optimal Protein Production **5.3.1**
Basic Protocol 1: Production of Recombinant Proteins in Batch Fermentations 5.3.1
Alternate Protocol 1: Protein Production in High-Cell-Density Processes: Fed-Batch Methods 5.3.8
Alternate Protocol 2: In Vivo Labeling of Recombinant Proteins with Heavy Metal Derivatives 5.3.9
Alternate Protocol 3: Stable Isotopic Labeling of Recombinant Proteins 5.3.10
Support Protocol 1: Growth Monitoring 5.3.11
Support Protocol 2: Sterility Checking 5.3.12
Reagents and Solutions 5.3.12
Commentary 5.3.14

5.4 Overview of the Baculovirus Expression System **5.4.1**
Baculovirus Life Cycle 5.4.1
Baculovirus Expression System 5.4.1
Post-Translation Modification of Proteins in Insect Cells 5.4.3
Literature Cited 5.4.3

continued

BCU IAST / CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

5.5 Protein Expression in the Baculovirus System 5.5.1
Basic Protocol 1: Large-Scale Production of Viral Stock 5.5.2
Basic Protocol 2: Determination of Expression Kinetics 5.5.3
Basic Protocol 3: Production in Bioreactors 5.5.5
Alternate Protocol 1: Production in Perfusion Cultures 5.5.6
Support Protocol 1: Titrating Viral Stocks with the Plaque Assay 5.5.8
Support Protocol 2: Titrating Viral Stocks with the End-Point Assay 5.5.9
Support Protocol 3: Harvesting 5.5.10
Support Protocol 4: Preparing Samples for Protein Analysis 5.5.11
Support Protocol 5: Monitoring Nutrient Levels in Culture Media 5.5.12
Support Protocol 6: Determination of Cell Growth and Viability 5.5.12
Reagents and Solutions 5.5.13
Commentary 5.5.13

5.6 Overview of Protein Expression in *Saccharomyces cerevisiae* 5.6.1
Vectors and Host Strains 5.6.2
Challenges to the Expression of Foreign Proteins 5.6.3
Time Frame and Future Prospects 5.6.6
Literature Cited 5.6.6

5.7 Overview of Protein Expression in *Pichia pastoris* 5.7.1
General Characteristics of *Pichia pastoris* 5.7.1
Heterologous Protein Expression 5.7.3
Commentary 5.7.13
Literature Cited 5.7.14

5.8 Culture of Yeast for the Production of Heterologous Proteins 5.8.1
Basic Protocol 1: Small-Scale Expression Using *S. cerevisiae* Galactose-Regulated Vectors 5.8.1
Alternate Protocol 1: Expression Using Glucose-Repressible *ADH2* Vectors 5.8.3
Alternate Protocol 2: Expression Using Vectors with Glycolytic Gene Promoters 5.8.4
Basic Protocol 2: Small-Scale Expression in *Pichia pastoris* 5.8.4
Basic Protocol 3: Large-Scale Expression Using *S. cerevisiae* Galactose-Regulated Vectors 5.8.6
Basic Protocol 4: Large-Scale Expression Using *Pichia pastoris* 5.8.8
Support Protocol: Small-Scale Preparation of Protein Extracts 5.8.10
Reagents and Solutions 5.8.11
Commentary 5.8.13

5.9 Overview of Protein Expression by Mammalian Cells 5.9.1
Choice of Mammalian Cell Host 5.9.1
Transfection, Selection, and Amplification 5.9.5
Protein Translation, Quality Control, and Covalent Modification 5.9.6
Growth of Mammalian Cells for Protein Expression 5.9.10
Scale of Operation 5.9.13
Summary 5.9.14
Literature Cited 5.9.15

BCU IASI / CENTRAL UNIVERSITY LIBRARY

5.10 Production of Recombinant Proteins in Mammalian Cells 5.10.1
Basic Protocol 1: Plasmid Purification by Alkaline-Lysis/Anion-Exchange Capture 5.10.3
Basic Protocol 2: Transfection of Cells by Lipofection 5.10.5
Alternate Protocol 1: Transfection of Cells by Electroporation 5.10.6
Basic Protocol 3: Selection for Neomycin Phosphotransferase (NPTII) with Geneticin (G418) 5.10.7
Support Protocol 1: Determination of Background Sensitivity to G418 5.10.8
Alternate Protocol 2: Screening of Green Fluorescent Protein by Flow Cytometry 5.10.8
Support Protocol 2: Single-Cell Cloning 5.10.9
Basic Protocol 4: Amplification of Dihydrofolate Reductase (DHFR) with Methotrexate (MTX) 5.10.10
Basic Protocol 5: Adaptation of Suspension Cells to Production Medium Through Multiple Passaging 5.10.12
Basic Protocol 6: Growth of Cells in Batch Mode in Large-Scale Spinner Culture Vessels 5.10.14
Basic Protocol 7: Growth of Cells in Large-Scale Batch Reactors 5.10.17
Alternate Protocol 3: Growth of Cells in Continuous Culture in Bioreactors 5.10.19
Harvesting of Protein Product from Batch Mode Spinner Cultures and Large-Scale Batch Reactors 5.10.20
Basic Protocol 8: Harvesting a Secreted Product 5.10.21
Basic Protocol 9: Harvesting a Cell-Associated Product 5.10.22
Support Protocol 3: Freezing Cells for Production of Cell Banks 5.10.23
Support Protocol 4: Numerical Determination of Specific Growth Rate of Cells 5.10.24
Support Protocol 5: Graphical Determination of Specific Growth Rate of Cells 5.10.24
Support Protocol 6: Numerical Determination of Specific Productivity of Cells Producing Heterologous Protein 5.10.25
Support Protocol 7: Graphical Determination of Specific Productivity of Cells Producing Heterologous Protein 5.10.25
Support Protocol 8: Preparation of Samples for Product Assay 5.10.26
Support Protocol 9: Determination of Shaker Flask Growth Curve 5.10.26
Reagents and Solutions 5.10.27
Commentary 5.10.28

BCU IASI // CENTRAL UNIVERSITY LIBRARY

CHAPTER 5
Production of Recombinant Proteins

INTRODUCTION

This chapter deals with the production of recombinant proteins using the expression systems in common use. Methods describing the actual construction of the expression vectors or plasmids containing the genes coding the proteins of interest are described in detail elsewhere (Sambrook et al., 1989; Ausubel et al., 1997).

The most widely used and most important expression systems are those using *E. coli* as the host cell. An overview of the production of proteins in *E. coli* is provided in *UNIT 5.1*. General strategies for gene expression described include direct intracellular expression, secretion of the protein into the periplasmic space, and the expression of fusion proteins. Some of these strategies (e.g., secretion procedures) are used to overcome the common problem of inclusion body formation (see *UNIT 6.1* for additional discussion); others (e.g., fusions involving affinity tags) are used simply to assist in protein purification. Many of the expression vectors discussed are commercially available and suppliers are indicated throughout the text.

Protocols describing protein expression in transformed *E. coli* cells are provided in *UNITS 5.2 & 5.3*. First, protein expression in *E. coli* requires that a suitable host strain (tabulated in *UNIT 5.2*) be transformed with an expression vector. Electroporation, a quick and simple method for transforming *E. coli* host strains, is presented in *UNIT 5.2*. Second, the growth conditions and methods used to induce gene expression are optimized for the particular host strain and controllable transcriptional promoter used. Standard procedures for small-scale expression in shaker flasks using several of the most common expression vectors are also presented, along with protocols for determining plasmid stability and storing transformed strains at low temperature. Finally, methods for establishing the level of protein expression as well as solubility and location (intracellular or periplasmic) are described.

Once suitable host/vector combinations have been selected on the basis of the small-scale experiments described in *UNIT 5.2*, large-scale protein production can be carried out in fermentors with careful monitoring of cell growth. *UNIT 5.3* describes fermentation in batch mode using a fixed amount of medium. For higher-density fermentations, the fed-batch process is described: in this procedure the growth carbon source is continuously fed into the culture. Modifications to the standard fed-batch process are used to introduce the heavy atom derivative selenomethionine into recombinant proteins to facilitate X-ray structure determination and to uniformly label proteins with combinations of the stable isotopes ^{2}H, ^{13}C, and ^{15}N. Isotopically labeled proteins are required for high-resolution structural studies using nuclear magnetic resonance (NMR) spectroscopy (see *UNIT 7.1* for additional details). Protocols for growth monitoring and sterility checking are also included.

There are several eukaryotic cell–based expression systems available for overproducing recombinant proteins (consult Ausubel et al., 1997, Chapter 16), and most of these will eventually be covered in this chapter. One system, which uses insect cells infected with recombinant baculovirus, has become widely used. An overview of protein expression using the baculovirus system is given in *UNIT 5.4*. References covering the molecular

Contributed by Paul T. Wingfield
Current Protocols in Protein Science (1998) 5.0.1-5.0.3
Copyright © 1998 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY LIBRARY

biology required to construct the recombinant baculovirus vector that expresses the gene of interest are provided in the overview. Furthermore, references to protocols describing small-scale expression experiments are supplied in the introduction to *UNIT 5.5*.

The large-scale growth of insect cells infected with baculovirus is carried out in bioreactors, requiring large (liter) amounts of viral stock (*UNIT 5.5*). These stocks must have known titers and are analyzed in small-scale (1- to 3-liter) cultures in order to evaluate and optimize protein expression. For developing protein purification methods and for isolating moderate amounts of recombinant protein, small-scale cultures in shaker or spinner culture flasks may provide adequate starting material. Large-scale insect cell culture is carried out with more sophisticated equipment using bioreactors in either batch mode or perfusion mode. Cell harvesting is carried out by centrifugation or tangential-flow microfiltration. The cells (for intracellular expressed proteins) or the culture medium (for secreted proteins) are retained for downstream processing. Other protocols in this unit provide methods for preparing samples for SDS-PAGE to check protein expression levels and for monitoring nutrient levels in culture media.

Yeast is another important eukaryotic host used for protein expression. The molecular biology of *Saccharomyces cerevisiae* (baker's yeast) is described in Chapter 13 of Ausubel et al. (1997). An overview of protein expression in *S. cerevisiae* is given in *UNIT 5.6*, which details, as a case study, some of the machinations required to achieve high-level production of the human glycoprotein antithrombin 111 secreted into the extracellular medium.

Other strains of yeast apart from *S. cerevisiae* have also been used for heterologous protein expression, including the methylotropic yeast *Pichia pastoris*. There may be some advantages to using this host system over the more common *S. cerevisiae*, as expression levels, especially for intracellular proteins, can be higher. Protein expression using the *P. pastoris* system is reviewed in *UNIT 5.7*. The commonly used expression vectors are described, together with many examples of proteins successfully expressed with the system. Protocols describing the culture of the *S. cerevisiae* and *P. pastoris* strains with various expression vectors for heterologous protein expression are provided in *UNIT 5.8*.

An overview of protein expression using mammalian cells is presented in *UNIT 5.9*. The unit reviews the stages involved in protein production using the stable expression approach. As post-transcriptional modification and intracellular protein transportation are integral components of mammalian protein expression, descriptions of these mechanisms are included. Protocols for the production of recombinant proteins in mammalian cells using various culture configurations are described in *UNIT 5.10*. These are supplemented with a number of support protocols that include the first steps in downstream processing.

The choice of protein expression system does depend on the intended use and characteristics of the protein. If a protein can be expressed in *E. coli* in a soluble state either directly into the cytoplasm or by secretion into the periplasm, then it may be possible to isolate large amounts of correctly folded and active protein (e.g., see *UNIT 6.2*). On the other hand, if the protein must be post-translationally modified (i.e., glycosylated or phosphorylated), a eukaryotic expression system must be used. Furthermore, if the protein is expressed in an insoluble state in *E. coli*, one way to circumvent this problem is to try expression in an eukaryotic system.

The baculovirus system complements *E. coli* expression systems and for many investigators is the system of choice, especially for expressing post-translationally modified proteins. Yeast has the advantage that it is cheap to grow (analogous to *E. coli*) but can produce hyperglycosylated proteins; moreover, occasionally protein produced at high levels is insoluble (again analogous to *E. coli*). If a glycoprotein is required for clinical

BCU IASI CENTRAL UNIVERSITY LIBRARY

studies, it must usually be produced by mammalian cell expression, as discussed in detail in *UNITS 5.9 & 5.10*. However, no expression system is without its drawbacks; some of the problems reported using the baculovirus and yeast systems are mentioned in *UNITS 5.4 & 5.6-5.8* and elsewhere (Bradley, 1990). In addition, a brief summary of the advantages and disadvantages of all the commonly used host systems is given in the booklet *The Recombinant Protein Handbook* (Pharmacia Biotech, 1994).

LITERATURE CITED

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds). 1997. *In* Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Bradley, M.K. 1990. Overexpression of proteins in eukaryotes. *Methods Enzymol.* 182:112-143.

Pharmacia Biotech. 1994. The Recombinant Protein Handbook (18-1105-02). Pharmacia Biotech, Piscataway, N.J.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular Cloning. A Laboratory Manual (2nd ed.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Paul T. Wingfield

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Production of Recombinant Proteins in *Escherichia coli*

The genetics and biochemistry of *Escherichia coli* are probably the best understood of any known organism. The knowledge gained in the study of *E. coli* biology has been applied to the development of many of today's molecular cloning techniques. Most cloning vectors and methods utilize *E. coli* or its phages as a preferred host, primarily because of the ease with which the bacterium can be grown and genetically manipulated. These same characteristics made *E. coli* an attractive early choice as a host for the production of large quantities of protein encoded by cloned genes. Aside from its well-studied biology, *E. coli* is suitable as the basis of an expression system because of its rapid doubling time and its ability to grow in inexpensive media. Years of study devoted to gene expression in *E. coli* have provided numerous choices for transcriptional and translational control elements that can be applied to the expression of foreign genes (*UNITS 5.2 & 5.3*). As a result, *E. coli* has been and continues to be the expression system of choice and a substantial body of literature has accumulated on the successful expression of foreign genes in this host. Several problems with protein expression in *E. coli* have been encountered, and many have been ultimately solved. This unit describes methods that have been developed for production of recombinant proteins in *E. coli* and potential pitfalls that may be encountered.

GENERAL STRATEGIES FOR GENE EXPRESSION IN *E. COLI*

The basic approach used to express foreign genes in *E. coli* begins with insertion of the gene into an expression vector, usually a plasmid (obtained from a commercial supplier or from the author of a published study). The next step involves transforming a suitable *E. coli* host strain with the plasmid, for example, by electroporation (*UNIT 5.2*). Transformed cells are then subjected to evaluation of plasmid stability and foreign protein expression following induction of the controllable transcriptional promoter present in the expression vector (*UNIT 5.2*). Once small-scale shaker flask experiments have identified successful expression systems (*UNIT 5.2*), the transformed *E. coli* strain can be used in large-scale fermentation systems (*UNIT 5.3*), which can be designed to produce recombinant proteins with altered amino acids (e.g., selenomethionine in place of methionine) or stable isotope tags to facilitate structural studies. Production is followed by protein purification (Chapter 6) and characterization (Chapter 7).

If the gene to be expressed in *E. coli* is of eukaryotic origin, it must be a cDNA copy, as *E. coli* will not recognize and splice out introns from transcripts of genomic copies of eukaryotic genes. Typical *E. coli* expression vectors contain the following elements.

Selectable Marker

Expression plasmids contain sequences encoding a selectable marker to ensure maintenance of the vector in the host cell. Commonly used selectable markers in *E. coli* include *bla* (which encodes β-lactamase and confers resistance to ampicillin and other β-lactam antibiotics), *cat* (which encodes chloramphenicol acetyltransferase and confers resistance to chloramphenicol), and *tet* (which encodes a membrane protein that confers resistance to tetracycline).

Origin of Replication

Replication of a plasmid as an independent extrachromosomal element is controlled by its origin of replication. Some plasmids are under so-called stringent control, wherein the plasmid's replication is coupled to replication of the host chromosome. As such, only one or at most a few copies of the plasmid are maintained in each cell. More commonly, *E. coli* vectors used for expression of cloned genes utilize relaxed control such as the ColE1 replicon found in pBR322 and derivatives. These plasmids generally have copy numbers of 10 to 200 plasmids per cell. The presence of multiple copies of the gene and its associated control elements can provide advantages in mRNA production and, as a result, may be reflected in an increase in the total amount of protein that accumulates. Nevertheless, the reduced copy number of stringent replicons may sometimes be advantageous: it allows for modulation of gene dosage, which can sometimes be helpful for altering the kinetics of expression or reducing the basal transcriptional levels of the plasmid-encoded promoter.

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

Promoters

An essential component of expression vectors is a controllable transcriptional promoter which, when induced, can direct the production of large amounts of mRNA from the cloned gene. There are a variety of controllable promoters that are routinely used. The *lac* promoter system utilizes the transcriptional control elements from the *E. coli* β-galactosidase gene (Miller and Reznikoff, 1978). The *lac* promoter (like the tac and trc promoters, which have optimized *lac* promoter RNA polymerase recognition sequences) is controlled by binding of *lac* repressor (the *lacI* gene product) to an operator sequence in the promoter region. The repressor can be expressed either from the host genome (single copy, in which case an overexpressing repressor allele called *lacI*^{q} should be used) or from the expression vector (multiple copies, with resultant tighter transcriptional control). Induction of the promoter is generally accomplished by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG), a lactose analog that binds to the *lac* repressor and prohibits its binding to the *lac* operator. An example of a commercially available *lac* promoter vector is pBluescript (Stratagene). The pKK223-3 vector from Pharmacia Biotech is a source of the tac promoter. The trc promoter is used in the vectors pSE280, pSE380, and pSE420 from Invitrogen.

Another commonly used promoter system utilizes the major leftward promoter of bacteriophage lambda, p_L, and its control elements (Shimatake and Rosenberg, 1981). The p_L promoter contained on an expression vector can be controlled by the phage-encoded *c*I repressor, which is typically expressed from an integrated copy of the phage in the host genome. A temperature-sensitive *c*I repressor (called *c*I857) is usually used; this encodes a repressor that is functional at lower temperatures but denatures at temperatures above 37.5°C. Thus, p_L-mediated protein synthesis can be induced by a simple temperature shift. Alternatively, temperature-independent p_L promoter systems have been developed that utilize a *c*I repressor gene under the control of a separate inducible promoter (Mieschendahl et al., 1986; LaVallie et al., 1993a). A commercial source for the p_L vector is the pP$_L$-Lambda vector from Pharmacia Biotech.

The T7 RNA polymerase promoter is another popular transcriptional control element for heterologous expression. The RNA polymerase from bacteriophage T7 is highly selective for specific T7 phage promoter sequences that are uncommon in other DNAs (Studier et al., 1990). A gene of interest placed under the control of a T7 promoter can be selectively expressed by induction of host-encoded T7 polymerase synthesis, which itself is under the control of an inducible promoter such as *lac*. Induction of T7 polymerase synthesis causes almost exclusive expression of the gene under the control of the T7 promoter. In some instances, the T7 transcription can outcompete the host RNA polymerase, resulting in accumulation of large amounts of the gene product of interest. There are many commercially available expression vectors that utilize the T7 polymerase/promoter system, such as the pGEMEX vectors from Promega, the pRSET vectors from Invitrogen, and the pET vectors from Novagen.

Translation Initiation Sequence

Initiation of translation on mRNAs requires the presence of a so-called Shine and Dalgarno sequence or ribosome binding site (RBS) in close proximity to an initiator methionine (Shine and Dalgarno, 1974). The RBS consists of a purine-rich stretch of nucleotides complementary to the 3′ end of 16S RNA, located 5 to 13 bases 5′ to an initiator ATG. RBS elements typically used in expression vectors derive from well-translated *E. coli* or bacteriophage genes. For instance, the pTrcHis and pRSET vectors (Invitrogen) and the pGEMEX (Promega) vectors use the T7 gene 10 RBS.

SPECIFIC EXPRESSION STRATEGIES

Direct Intracellular Expression

Direct expression refers to the fusing of the coding sequence of interest to transcriptional and translational control sequences on an expression vector, with an initiator methionine codon preceding the open reading frame. This approach can be used to produce cytoplasmic proteins, and it can also be used for the intracellular expression of normally secreted proteins. In the latter case, the DNA sequence encoding the signal peptide is replaced by the initiator methionine codon.

Success with the direct approach is often variable. First, translation initiation is inconsistent due to the fact that sequences 3′ to the initiator methionine can influence the efficiency of ribosome binding (Looman et al., 1987; Bucheler et al., 1990). For reasons that are not fully understood, maximizing the A+T content of the 5′ end of the coding sequence (taking advantage of the degeneracy of the

genetic code) can sometimes improve the efficiency of translation initiation (De Lamarter et al., 1985; Devlin et al., 1988). Second, recombinant proteins produced in the cytoplasm often form dense, insoluble aggregates of protein called *inclusion bodies* (Schein, 1989). In some ways, this can be viewed as an advantage because inclusion bodies are easily purified from the soluble proteins (*UNIT 6.3*). In addition, proteins in inclusion bodies are usually resistant to proteolytic degradation. In cases where activity of the expressed protein is unnecessary (e.g., in production of protein to be used as an antigen to produce antibodies), inclusion body formation and resultant insolubility of the protein may actually be preferred. If active protein is desired, however, recovery of properly folded protein from inclusion bodies requires the total denaturation of the protein using reagents such as urea or guanidine, followed by subsequent refolding using protocols that must be determined empirically for each protein (*UNIT 6.5*). The success of refolding protocols is variable, depending on the particular protein, and may yield little or no correctly folded material.

If protein insolubility is encountered and is undesired, production of soluble protein can sometimes be enhanced by simply lowering the growth temperature during protein synthesis (Bishai et al., 1987; Schein and Noteborn, 1988). If the protein is normally secreted, then a secretory construct may produce more soluble protein (see section on Secretion). Alternatively, production of the protein as a fusion to a highly soluble partner may alleviate the problem (see section on Fusion Proteins).

If the gene is expressed poorly, translation efficiency may be a problem. If this is the case, it is often advantageous to maximize the A+T content of the 5′ coding sequence and/or replace the RBS with a sequence that initiates translation more efficiently (Olins and Rangwala, 1989). Another potential cause of poor translation is a preponderance of "rare" codons in the coding sequence, that is, codons that are underutilized by *E. coli* (Robinson et al., 1984). Stretches of contiguous rare codons may contribute to lower expression levels and should be changed to codons that are frequently used by *E. coli*. Finally, poor expression levels may be caused by product instability. In this instance, use of host strains and induction methods that minimize proteolytic degradation should be attempted (*UNIT 5.2*).

In instances where the gene is expressed well and soluble protein can be produced, the protein must then be purified from an abundant and diverse mixture of other *E. coli* cytoplasmic proteins. This can be a difficult and time-consuming process, and the purification scheme will be different for each protein based on its physical and biochemical characteristics (Chapters 6 and 8).

Secretion

Secretion of proteins in *E. coli* is mediated by the presence of an N-terminal signal sequence that is cleaved after translocation of the protein. Expression of cloned gene products as secreted proteins in *E. coli* has been utilized as an alternative to cytoplasmic expression for proteins that are normally secreted. In *E. coli* the protein is secreted to the periplasmic space between the cytoplasmic and outer membranes, in contrast to extracellular secretion that occurs in gram-positive bacteria and eukaryotic cells. The result is that in *E. coli* the secreted protein remains cell-associated, although in a "compartment" separated from the cytoplasmic proteins that make up the vast majority of the total cellular protein. This can be advantageous in terms of protein purification if techniques are used that release only periplasmic contents while leaving the cytoplasmic membrane intact (Neu and Heppel, 1965).

Secretion of heterologous gene products has been successfully employed for various proteins that are difficult to produce in the cytoplasm of *E. coli* as soluble and active proteins, including various growth factors (Cheah et al., 1994), receptors (Fuh et al., 1990), and recombinant Fab fragments (Skerra, 1994). Although eukaryotic signal peptides have been reported to function in *E. coli*, most secretion vectors utilize signal peptides derived from prokaryotic genes such as the *ompA* (Takahara et al., 1988; Cheah et al., 1994), *pelB* (Power et al., 1992), *phoA* (Oka et al., 1985), or *hisJ* (Vasquez et al., 1989) signal peptides. Although secretion provides an alternative to cytoplasmic expression that can sometimes result in the production of properly folded and active protein, the yield of desired protein is often low. Also, overexpressed gene products have been reported to form inclusion bodies even in the periplasm (Bowden and Georgiou, 1990); so secretion is not a panacea for the problem of insolubility.

Fusion Proteins

The problems that have plagued successful overexpression of cloned gene products in *E. coli*—namely, inconsistent expression levels,

BRUNEL UNIVERSITY LIBRARY

protein insolubility, and difficult purification of the gene product from *E. coli* contaminants—have been most successfully addressed by the use of fusion proteins. Fusion proteins are created via a translational fusion of the coding sequence for the protein of interest to a gene for a highly expressed protein partner (or "carrier" protein). Typically, the gene encoding the protein of interest is inserted in-frame 3′ to the coding sequence for the carrier protein, in place of the usual termination codon. This allows uniform translational initiation of the carrier protein regardless of the coding sequence fused to its 3′ end, which helps to ensure consistent expression levels.

Carrier proteins are usually chosen based on specific attributes that make them suitable in this role. The most successful fusion systems employ the maltose-binding protein (MBP; Maina et al., 1988), glutathione *S*-transferase (GST; Smith and Johnson, 1988), or thioredoxin (TRX; LaVallie et al., 1993a). The genes for these proteins are well expressed, and the proteins are highly soluble and provide specific physical characteristics or affinities to aid purification. These qualities also extend to the protein sequences fused to them, thereby enhancing the solubility and ease of purification of the entire fusion protein.

Maltose-binding protein is a 43-kDa secreted protein from *E. coli*. As its name implies, it binds specifically to maltose or amylose, a property that can be exploited in purification schemes (Guan et al., 1988). MBP fusions can be secreted into the periplasm, or they can be expressed without the MBP signal peptide, which results in accumulation of the fusion protein in the cytoplasm. MBP fusions are usually well expressed, and a high proportion of MBP fusion proteins are soluble. MBP expression plasmids and reagents can be purchased from New England Biolabs. The MBP expression vectors utilize the tac promoter and a plasmid-encoded *lacI*^{q} gene for transcriptional control. These vectors contain a recognition sequence for the site-specific protease factor Xa to allow removal of the carrier protein following purification (see section on Fusion Protein Cleavage Methods).

Glutathione *S*-transferase is a 26-kDa cytoplasmic protein from *Schistosoma japonicum*. Carboxy-terminal protein fusions to GST are usually soluble and well-expressed (Smith and Johnson, 1988). GST binds specifically to glutathione, and GST fusion proteins can be purified in a single step from crude bacterial lysates by affinity chromatography on immobilized glutathione. The GST expression plasmids are available from Pharmacia Biotech, and utilize the tac promoter and plasmid-encoded *lacI*^{q} for inducible transcriptional control. A variety of vectors are available which contain either the factor Xa or the thrombin recognition sequence to allow cleavage and removal of the GST from the protein of interest following affinity purification (see section on Fusion Protein Cleavage Methods).

Thioredoxin is a 12-kDa intracellular *E. coli* protein. It is very soluble, can be highly overexpressed from plasmid vectors (Lunn et al., 1984), and has been shown to be a very successful fusion partner. A wide variety of gene products can be produced abundantly in soluble fashion when fused to TRX (LaVallie et al., 1993a). Two different characteristics of TRX can be exploited to allow specific purification of some TRX fusion proteins. TRX accumulates at specific sites along the inner surface of the cytoplasmic membrane, called Bayer's patches or adhesion zones. These sites constitute an osmotically sensitive compartment, and TRX (and some TRX fusions) can be selectively released from the cytoplasm and separated from the bulk of *E. coli* proteins by osmotic shock or freeze/thaw procedures. In addition, TRX is thermostable, and some TRX fusions can be purified by selective thermal denaturation of contaminants. In instances where osmotic shock or heat treatments do not provide adequate purification, an altered TRX protein has been developed (E.R.L., manuscript in preparation) that allows specific purification by metal-chelate affinity chromatography. The TRX fusion vector (available from Invitrogen) contains an enterokinase recognition sequence positioned at the junction between TRX and the C-terminal fusion partner.

In addition to the carrier proteins listed above, there are many other proteins that have been used to produce fusions for the purpose of generating large amounts of protein in *E. coli*. Among them are *E. coli* β-galactosidase (Ruther and Muller-Hill, 1983) and TrpE (Yansura, 1990), *Staphylococcus aureus* protein A (Nilsson et al., 1985; vector available from Pharmacia Biotech), chloramphenicol acetyltransferase (Knott et al., 1988), bacteriophage lambda cII protein (Nagai and Thøgersen, 1987), and various carbohydrate-binding proteins (Taylor and Drickamer, 1991; Helman and Mantsala, 1992). Table 5.1.1 lists many fusion partners that have been described in the literature.

ECU IASTATE UNIVERSITY LIBRARY

Table 5.1.1 Common Fusion Proteins

Fusion protein	Specific purification method[a]
Maltose-binding protein	Amylose binding
Glutathione *S*-transferase	Glutathione binding
Thioredoxin	Selective release, heat stability, MCAC
Protein A	IgG binding
β-Galactosidase	APTG or anti-β-galactosidase antibody binding
Chloramphenicol acetyltransferase	Chloramphenicol binding
lac repressor	*lac* operator binding (Lundeberg et al., 1990)
Galactose-binding protein	Galactose binding
Cyclomaltodextrin glucanotransferase	Cyclodextrin binding
Lambda *c*II protein	None
TrpE protein	None

[a]MCAC, metal-chelate affinity chromatography; IgG, immunoglobulin G; APTG, *p*-amino-β-D-thiogalactoside.

Table 5.1.2 Fusion Tags

Amino acid tag	Ligand	Reference/supplier
Polyhistidine	Metal-chelate affinity resin	Qiagen
Polyaspartic acid	Anion-exchange resin	Dalbøge et al. (1987)
Polyarginine	Cation-exchange resin	Brewer and Sassenfeld (1985)
Polyphenylalanine	HIC resin[a]	Persson et al. (1988)
Polycysteine	Thiol	Persson et al. (1988)
In vivo biotinylated peptide	Avidin or streptavidin	Schatz (1993)
Flag peptide	Anti-Flag antibody	International Biotechnologies (IBI)

[a]HIC, hydrophobic interaction chromatography (*UNIT 8.4*).

Fusion Tags

Fusion tags are small stretches of amino acids added to the N-terminal or C-terminal end of a protein. Although they do not usually help to increase expression levels or protein solubility, they can be advantageous in protein purification and detection. The tags are generally chosen either because they encode an epitope that can be detected and purified using an antibody that binds to it, or because the tag amino acids provide a physical characteristic that can be exploited for easy and specific purification. A popular example is the polyhistidine tag, usually a stretch of six consecutive histidine residues added to either the N or C terminus of a protein, that provides specific binding to metal chelate resins. There are a number of polyhistidine vectors on the market, such as the pTrcHis vector from Invitrogen. Other examples, such as polyarginine (Brewer and Sassenfeld, 1985) or polyaspartic acid tags (Dalbøge et al., 1987), can be used to alter the binding behavior of a protein on ion-exchange resins. A noteworthy fusion tag is a small stretch of amino acids recognized and biotinylated in vivo by the biotin-protein ligase in *E. coli* (Schatz, 1993). This allows specific capture of the biotinylated protein on immobilized avidin or streptavidin. Table 5.1.2 lists a number of the more popular fusion tags.

Fusion Protein Cleavage Methods

Whereas fusion proteins and fusion tags have proved useful for the reasons outlined above, the end result is that the protein of interest will carry additional sequences that may hamper its functional activity. There are several methods for removal of the carrier protein, which can be divided into chemical cleavage methods and enzymatic (proteolytic) cleavage methods. Chemical cleavage of fusion proteins can be accomplished with reagents such as cyanogen bromide (Met↓; Itakura et al., 1977; Villa et al., 1989), 2-(2-

BOSTON CENTRAL UNIVERSITY LIBRARY

nitrophenylsulfenyl)-3-methyl-3′-bromoindolinine or BNPS-skatole (Trp↓; Dykes et al., 1988), hydroxylamine (Asn↓Gly; Bornstein and Balian, 1977), or low pH (Asp↓Pro; Szoka et al., 1986). The use of these reagents, as described in *UNIT 11.4*, can be applied to the site-specific cleavage of fusion proteins provided that the fusion is designed so that the chemically labile bond is positioned at the desired point of scission. Chemical cleavage reagents tend to be inexpensive and efficient, and many of the reactions can be performed under denaturing conditions so that even insoluble proteins can be cleaved. However, one disadvantage of using chemical cleavage reagents for the site-specific cleavage of fusion proteins is that the reactions are generally performed under extremes of pH and/or temperature that can result in unwanted amino acid side-chain modifications. Chemical cleavage methods also have the disadvantage of low specificity, so it is more likely that the protein of interest will contain an internal cleavage site.

Enzymatic digestion is usually the method of choice for soluble fusion protein cleavage as the reactions are carried out under relatively mild conditions. Proteases such as trypsin (Lys↓ or Arg↓), endoproteinase Asp-N (↓Asp), and Staph V8 *S. aureus* V8 protease (Glu↓ or Asp↓) can be used in this role. High-quality trypsin, *S. aureus* V8 protease (endoproteinase Glu-C), and endoproteinase Asp-N can all be obtained from Boehringer Mannheim. These proteases, like chemical cleaving reagents, are limited by their low degree of specificity (recognizing and cleaving at single amino acids).

Alternatively, thrombin (Leu-Val-Pro-Arg↓Gly-Ser; Gearing et al., 1989), factor Xa [Ile-Glu(or Asp)-Gly-Arg↓; Nagai and Thøgersen, 1984, 1987; Gardella et al., 1990], renin (Pro-Phe-His-Leu↓Leu-Val-Tyr; Haffey et al., 1987), collagenase (Pro-X↓Gly-Pro-Y↓; Germino and Bastia, 1984), and enterokinase (Asp-Asp-Asp-Asp-Lys↓; Dykes et al., 1988; LaVallie et al., 1993b) are much more suitable in this role. All of these enzymes have extended substrate recognition sequences (up to seven amino acids in the case of renin), which greatly reduces the likelihood of unwanted cleavages elsewhere in the protein. Factor Xa and enterokinase are most useful for cleaving off C-terminal fusion partners because they cleave on the carboxyl-terminal side of their respective recognition sequences. This allows the release of fusion partners containing their authentic amino terminus. In addition, the catalytic subunit of bovine enterokinase has been cloned and expressed (LaVallie et al., 1993b), and the recombinant enzyme has been shown to be approximately 100-fold more efficient than the native intestinally derived enzyme in fusion protein cleavage reactions (Racie et al., 1995). The recombinant enterokinase is available from New England Biolabs.

SUMMARY

Methods for the overexpression of cloned gene products in *E. coli* have improved significantly since it was first attempted. Common problems such as variable expression levels, inclusion body formation, and purification difficulties have been successfully addressed by advancements in expression technology. Probably the most significant of these advancements has been the development of fusion proteins and fusion tag expression and purification techniques. These methods have resulted in more consistent production of soluble and active protein, and have allowed for simple and efficient purification of the proteins from bacterial lysates. Although the production of soluble, properly folded, and active recombinant proteins in *E. coli* is still not guaranteed, the likelihood of success is far greater than it was just a few years ago. This progress should help ensure that *E. coli* will continue to be the host organism of choice for recombinant protein production.

LITERATURE CITED

Bishai, W.R., Rappuoli, R., and Murphy, J.R. 1987. High-level expression of a proteolytically sensitive diphtheria toxin fragment in *Escherichia coli*. *J. Bacteriol.* 169:5140-5151.

Bornstein, P. and Balian, G. 1977. Cleavage at Asn-Gly bonds with hydroxylamine. *Methods Enzymol.* 47:132-145.

Bowden, G.A. and Georgiou, G. 1990. Folding and aggregation of β-lactamase in the periplasmic space of *Escherichia coli*. *J. Biol. Chem.* 265:16760-16766.

Brewer, S.J. and Sassenfeld, H.M. 1985. The purification of recombinant proteins using C-terminal poly-arginine fusions. *Trends Biotechnol.* 3:119-122.

Bucheler, U.S., Werner, D., and Schirmer, R.H. 1990. Random silent mutagenesis in the initial triplets of the coding region: A technique for adapting human glutathione reductase-encoding cDNA to expression in *Escherichia coli*. *Gene* 96:271-276.

Cheah, K.C., Harrison, S., King, R., Crocker, L., Well, J.R., and Robins, A. 1994. Secretion of eukaryotic growth hormones in *Escherichia coli* is influenced by the sequence of the mature proteins. *Gene* 138:9-15.

RSU IASTRE LIBRARY

Dalbøge, H., Dahl, H.H.M., Pedersen, J., Hansen, J.W., and Christensen, T. 1987. A novel enzymatic method for production of authentic hGH from an *Escherichia coli*-produced hGH precursor. *Bio/Technology* 5:161-164.

De Lamarter, J.F., Mermod, J.J., Liang, C.M., Eliason, J.F., and Thatcher, D.R. 1985. Recombinant murine GM-CSF from *E. coli* has biological activity and is neutralized by a specific antiserum. *EMBO J.* 4:2575-2581.

Devlin, P.E., Drummond, R.J., Toy, P., Mark, D.F., Watt, K.W.K., and Devlin, J.J. 1988. Alteration of the amino-terminal codons of human granulocyte colony-stimulating factor increases expression levels and allows efficient processing by methionine aminopeptidase in *Escherichia coli*. *Gene* 65:13-22.

Dykes, C.W., Bookless, A.B., Coomber, B.A., Noble, S.A., Humber, D.C., and Hobden, A.N. 1988. Expression of atrial natriuretic factor as a cleavable fusion protein with chloramphenicol acetyltransferase in *Escherichia coli*. *Eur. J. Biochem.* 174:411-416.

Fuh, G., Mulkerrin, M.G., Bass, S., McFarland, N., Brochier, M., Bourell, J.H., Light, D.R., and Wells, J.A. 1990. The human growth hormone receptor. Secretion from *Escherichia coli* and disulfide bonding pattern of the extracellular binding domain. *J. Biol. Chem.* 265:3111-3115.

Gardella, T.J., Rubin, D., Abou-Samra, A.-B., Keutmann, H.T., Potts, J.T. Jr., Kronenberg, H.M., and Nussbaum, S.R. 1990. Expression of human parathyroid hormone (1-84) in *Escherichia coli* as a factor X-cleavable fusion protein. *J. Biol. Chem.* 265:15854-15859.

Gearing, D.P., Nicola, N.A., Metcalf, D., Foote, S., Willson, T.A., Gough, N.M., and Williams, R.L. 1989. Production of leukemia factor in *Escherichia coli* by a novel procedure and its use in maintaining embryonic stem cells in culture. *Bio/Technology* 7:1157-1161.

Germino, J. and Bastia, D. 1984. Rapid purification of a gene product by genetic fusion and site specific proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* 81:4692-4696.

Guan, C., Li, P., Riggs, P.D., and Inouye, H. 1988. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67:21-30.

Haffey, M.L., Lehman, D., and Boger, J. 1987. Site-specific cleavage of a fusion protein by renin. *DNA* 6:565-571.

Hellman, J. and Mantsala, P. 1992. Construction of an *E. coli* export-affinity vector for expression and purification of foreign proteins by fusion to cyclomaltodextrin glucanotransferase. *J. Biotechnol.* 23:19-34.

Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* 198:1056-1063.

Knott, J.A., Sullivan, C.A., and Weston, A. 1988. The isolation and characterisation of human atrial natriuretic factor produced as a fusion protein in *Escherichia coli*. *Eur. J. Biochem.* 174:405-410.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993a. A thioredoxin gene fusion system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology* 11:187-193.

LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., and McCoy, J.M. 1993b. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268:23311-23317.

Looman, A.C., Bodlaender, J., Comstock, L.J., Eaton, D., Jhurani, P., de Boer, H.A., and van Knippenberg, P.H. 1987. Influence of the codon following the AUG initiation codon on the expression of a modified *lacZ* gene in *Escherichia coli*. *EMBO J.* 6:2489-2492.

Lundeberg, J., Wahlberg, J., and Uhlen, M. 1990. Affinity purification of specific DNA fragments using a *lac* repressor fusion protein. *Genet. Anal. Techn. Appl.* 7:47-52.

Lunn, C.A., Kathju, S., Wallace, B.J., Kushner, S.R., and Pigiet, V. 1984. Amplification and purification of plasmid-encoded thioredoxin from *Escherichia coli* K12. *J. Biol. Chem.* 259:10469-10474.

Maina, C.V., Riggs, P.D., Grandea, A.G., Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A., and Guan, C. 1988. An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene* 74:365-373.

Mieschendahl, M., Petri, T., and Hanggi, U. 1986. A novel prophage independent *trp* regulated lambda p_L expression system. *Bio/Technology* 4:802-808.

Miller, J.H. and Reznikoff, W.S. (eds.) 1978. The Operon. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Nagai, K. and Thøgersen, H. C. 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein in *Escherichia coli*. *Nature* 309:810-812.

Nagai, K. and Thøgersen, H.C. 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. *Methods Enzymol.* 153:461-481.

Neu, H.C. and Heppel, L.A. 1965. Release of enzymes from *Escherichia coli* by osmotic shock during the formation of spheroplasts. *J. Biol. Chem.* 240:3685-3692.

Nilsson, B., Abrahmsen, L., and Uhlen, M. 1985. Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors. *EMBO J.* 4:1075-1080.

Oka, T., Sakamoto, S., Miyoshi, K., Fuwa, T., Yoda, K., Yamasaki, M., Tamura, G., and Miyake, T. 1985. Synthesis and secretion of human epider-

BIBLIOTECA CENTRAL UNIVERSITY LIBRARY

mal growth factor by *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82:7212-7216.

Olins, P.O. and Rangwala, S.H. 1989. A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the *lacZ* gene in *Escherichia coli*. *J. Biol. Chem.* 264:16973-16976.

Persson, M., Bergstrand, M.G., Bulow, L., and Mosbach, K. 1988. Enzyme purification by genetically attached polycysteine and polyphenylalanine tags. *Anal. Biochem.* 172:330-337.

Power, B.E., Ivancic, N., Harley, V.R., Webster, R.G., Kortt, A.A., Irving, R.A., and Hudson, P.J. 1992. High-level temperature-induced synthesis of an antibody VH-domain in *Escherichia coli* using the PelB secretion signal. *Gene* 113:95-99.

Racie, L.A., McColgan, J.M., Grant, K.L., DiBlasio-Smith, E.A., McCoy, J.M., and LaVallie, E.R. 1995. Production of recombinant bovine enterokinase catalytic subunit in *Escherichia coli* using the novel secretory fusion partner DsbA. *Bio/Technology*. In press.

Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarranton, G., Stephens, P., Millican, A., Eaton, M., and Humphreys, G. 1984. Codon usage can affect efficiency of translation of genes in *Escherichia coli*. *Nucl. Acids Res.* 12:6663-6671.

Ruther, U. and Muller-Hill, B. 1983. Easy identification of cDNA clones. *EMBO J.* 2:1791-1794.

Schatz, P.J. 1993. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Bio/Technology* 11:1138-1143.

Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *Bio/Technology* 7:1141-1149.

Schein, C.S. and Noteborn, M.H.M. 1988. Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. *Bio/Technology* 6:291-294.

Shimatake, H. and Rosenberg, M. 1981. Purified λ regulatory protein cII positively activates promoters for lysogenic development. *Nature* 292:128-132.

Shine, J. and Dalgarno, L. 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U.S.A.* 71:1342-1346.

Skerra, A. 1994. A general vector, pASK84, for cloning, bacterial production, and single-step purification of antibody Fab fragments. *Gene* 141:79-84.

Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* 67:31-40.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185:60-89.

Szoka, P.R., Schreiber, A.B., Chan, H., and Murthy, J. 1986. A general method for retrieving components of a genetically engineered fusion protein. *DNA* 5:11-20.

Takahara, M., Sagai, H., Inouye, S., and Inouye, M. 1988. Secretion of human superoxide dismutase in *Escherichia coli*. *Bio/Technology* 6:195-198.

Taylor, M.E. and Drickamer, K. 1991. Carbohydrate-recognition domains as tools for rapid purification of recombinant eukaryotic proteins. *Biochem. J.* 274:575-580.

Vasquez, J.R., Evnin, L.B., Higaki, J.N., and Craik, C.S. 1989. An expression system for trypsin. *J. Cell. Biochem.* 39:265-276.

Villa, S., DeFazio, G., and Canosi, U. 1989. Cyanogen bromide cleavage at methionine residues of polypeptides containing disulfide bonds. *Anal. Biochem.* 177:161-164.

Yansura, D.G. 1990. Expression as *trpE* fusion. *Methods Enzymol.* 165:161-166.

Contributed by Edward R. LaVallie
Genetics Institute, Inc.
Cambridge, Massachusetts

PSU IASI / CENTRAL UNIVERSITY LIBRARY

Selection of *Escherichia coli* Expression Systems

Use of *Escherichia coli* as a protein factory capable of being retooled to meet the needs of the investigator is a well-established alternative to purification of proteins from natural sources. Moreover, expression—the directed synthesis of a foreign gene—is often the next step for researchers who have isolated a gene and want to study the protein it encodes.

Table 5.2.1 lists the most useful expression strains for fermentation processes. A more comprehensive table of *E. coli* strains can be found in Raleigh et al. (1989) and Sambrook et al. (1989). A summary of four of the most widely used expression systems is presented in Table 5.2.2. The available expression systems have been reviewed (e.g., Goeddel, 1990). A number of systems use thermoinducible promoters that predictably induce the *E. coli* heat-shock response. We therefore frequently use strains partially or totally defective in this response (e.g., Coli B, HB101, or DH5α) since some of the induced heat-shock proteins are proteases.

Table 5.2.1 *Escherichia coli* Fermentation Strains

Strain	Genotype	Source	Comment
B	Wild type	ATCC 23227/nonK12 lon deficient	Commonly used
B834	F⁻ *ompT gal met* r_Bm_B (DE3)[a]	Novagen	Host for selenomethionine labeling with T7 expression
BL21	F⁻ *ompT* r_Bm_B (DE3)[a]	Novagen	Host for T7 expression
DH1	F– *gyrA96 recA1 relA1 endA1 thi-1 hsdR17 supE44* λ⁻ *mcrA mcrB nalR*	ATCC 33849	General host
DH5α	F⁻ *gyrA96 recA1 relA1 endA1 thi-1 hsdR17 supE44* δ*lacU169* (ϕ*80lacZ*δ*M15*) λ⁻ *mcrA mcrB nalR*	GIBCO/BRL	General host
HMS174	F⁻*recA* r^-_{K12} m^+_{K12} Rif^R	Novagen	Host for T7 expression
HB101	F⁻ *leuB6 thi-1 hsdS20 lacY1 proA2 ara-14 galK2 xyl-5 mtl-1 recA13 rpsL20 supE44* λ⁻ *mcrA strR*	ATCC 33694	Inhibited at 42°C
JM101	*supE thi-1* δ(*lac-proAB*) F′[*traD36 proAB*⁺ $lacI^q$ *lacZ*δ*M15*]	ATCC 33876	Host for *lac* and *tac* expression
MM294	F⁻ *endA1 thi-1 hsdR17 supE44* λ⁻ *mcrA mcrB*	ATCC 33625	General host
RV308	*su*⁻ Δ*lac*X74 *gal*IISII::OP*308 strA*	ATCC 31608	Host for secretion
SG936	F⁻ *lac*(*am*) *trp*(*am*) *pho*(*am*) *supC*(*ts*) *rpsL mal*(*am*) *htpR*(*am*) *tsx*::Tn*10 lonR9*	ATCC 39624	Defective in heat-shock response
W3110	F⁻ λ⁻ *mcrA mcrB*	ATCC 27325	General host

[a]λDE3 lysogen

Copyright © 1995 by John Wiley & Sons, Inc.

BCU LIBRARY CENTRAL UNIVERSITY

Table 5.2.2 Summary of Expression Systems

Promoter	Repressor	Induction	Host	Selection	Reference
p_L	cI857	30°-42°C	*E. coli* B	Tet	Remaut et al., 1987
lac, tac	$lacI^Q$	IPTG	*E. coli* JM101	Amp	Stahl et al., 1989
trp	Trp	IAA[a]	*E. coli* B	Tet	Rosenberg et al., 1984
T7	$lacI^Q$	IPTG	*E. coli* BL21 (DE3) *E. coli* HMS174 (DE3)	Amp/Cm	Studier et al., 1986

[a]The *trp* promoter is naturally induced by tryptophan starvation; addition of IAA is optional.

This unit describes standard procedures for several expression systems, namely, temperature induction via the p_L promoter (Basic Protocol 1) and chemical induction via the *trp* promoter (Basic Protocol 2), *lac* or tac promoters (Alternate Protocol 1), and the T7 promoter (Alternate Protocol 2). These protocols require that the gene encoding the protein of interest has been identified and cloned into an appropriate expression vector using standard molecular biology techniques. Transformation of a suitable host strain (e.g., by electroporation; Support Protocol 2) is also a prerequisite.

Once an appropriate expression strain has been selected, the stability of the plasmid can be analyzed (Support Protocol 2) and the strain can be stored (Support Protocol 3). Other support protocols in this unit describe how to prepare samples for electrophoresis (Support Protocol 4), how to analyze the solubility of the expressed proteins (Support Protocol 5), and how to make samples of periplasmic extracts (Support Protocol 6) and extracellular media (using TCA precipitation, Support Protocol 7). Many of the support protocols are small-scale analysis procedures that are used to guide subsequent purification strategies and determine the suitability of the expression system for further development and scale-up.

BASIC PROTOCOL 1

EXPRESSION UNDER CONTROL OF p_L PROMOTER: TEMPERATURE INDUCTION

Protein expression from vectors containing the p_L promoter from bacteriophage lambda can be induced by raising the temperature. The *E. coli* host strains used with such vectors must harbor a temperature-sensitive repressor protein encoded by *cI857* either integrated in the host chromosome as a dormant phage, or built into the same plasmid as the one carrying the heterologous gene. At high temperatures, expression of the *cI857* gene is inactive, releasing the repression and allowing transcription of genes downstream of the p_L promoter. In this protocol, cells are first grown at 30°C to the desired density in rich medium containing antibiotic to maintain the plasmid; synthesis of the protein is then induced by a temperature shift to 42°C. The proper antibiotic is determined by the expression vector.

Materials

Host *E. coli* strain (e.g., Coli B, W3110) transformed with plasmid containing gene for the protein of interest under control of p_L promoter (see Support Protocol 1)
LB medium (see recipe)
Antibiotic stock solution (see recipe)

20-ml sterile culture tube
42°C shaking water bath
150-ml shaker flask, sterile

BGU IASI CENTRAL UNIVERSITY LIBRARY

1. Prepare a small liquid culture in stationary phase of the host *E. coli* strain transformed with plasmid containing the gene for the protein of interest under control of the p_L promoter by inoculating a 20-ml tube containing 2 ml LB medium (supplemented with appropriate antibiotic stock solution) with a single colony from a fresh agar plate of the same medium. Incubate the tube overnight on a rotary shaker at 30°C, 200 rpm. Do not allow the temperature to exceed 30°C.

   Temperatures above 30°C may induce expression of the target protein and interfere with growth.

2. The next morning, read the OD_{600} against water as a blank to verify that cultures have reached stationary phase.

   In a 1-cm light path cuvette, the OD should be in the range of 1 to 2. (An OD of 1 is equal to 0.5-1 × 10^9 cells/ml.) The OD is read on an aliquot of the culture diluted 10-fold in PBS buffer against water as a blank. Readings higher than 0.4 are not correlated to cell density. Absolute absorbance values will vary among spectrophotometers depending on the cuvettes, light path, and construction of the apparatus.

3. Place a 150-ml shaker flask containing 50 ml LB medium with antibiotic in a 42°C shaking water bath to prewarm the medium.

   Again, make sure that the temperature is accurate; too low a temperature will lead to incomplete repressor inactivation and poor induction. To measure temperature, insert a thermometer into a control flask (containing water only) also placed in the water bath.

4. Transfer the 2-ml overnight culture into the flask containing prewarmed medium. Take a 1-ml sample and process it as described in step 5. Return the 50-ml culture to the 42°C shaking water bath and incubate 3 hr with vigorous agitation (200 rpm). At the end of the incubation, take another 1-ml sample.

   Inoculation and sample removal should be done quickly to keep the temperature stable.

5. Microcentrifuge each sample 2 min at maximum speed, room temperature. Discard the supernatant and prepare for analysis by SDS-PAGE (see Support Protocol 4).

   Pellets may be stored at –20°C at this point before SDS-PAGE. They may be stored frozen indefinitely (for years).

6. Use the remainder of the 50-ml culture to analyze plasmid stability (Support Protocol 2), determine solubility of the sample (see Support Protocol 5), and prepare periplasmic extracts or extracellular medium samples (see Support Protocol 5 and Support Protocol 6) to obtain recombinant protein. Store leftover culture, if desired (see Support Protocol 3).

EXPRESSION UNDER CONTROL OF *trp* PROMOTER: CHEMICAL INDUCTION

BASIC PROTOCOL 2

This protocol describes induction of synthesis of the protein of interest under control of the *E. coli trp* promoter. Cells are grown to exponential phase in minimal medium to achieve tryptophan starvation, and residual *trp* repressor is inactivated by the tryptophan analog indole-3-acrylic acid. Incubation of cultures is conducted at 37°C throughout, and the choice of antibiotic is determined by the plasmid construction.

Materials

Host *E. coli* strain transformed with plasmid containing *trp* promoter and gene for the protein of interest (see Support Protocol 1)
M9 medium (see recipe) with 0.5% (w/v) Casamino Acids (Difco)
Antibiotic stock solution (see recipe)

BOU MASTER CENTRAL UNIVERSITY LIBRARY

Indole-3-acrylic acid (IAA; see recipe)

20-ml sterile culture tube
150-ml shaker flask
37°C orbital shaker

1. Use a single colony of the host *E. coli* strain transformed with plasmid containing *trp* promoter and gene for the protein of interest, from a fresh agar plate, to inoculate a 20-ml tube containing 2 ml of M9 medium with 0.5% Casamino Acids and appropriate antibiotic. Incubate the tube overnight at 37°C on a rotary shaker (200 rpm).
2. The next morning, read OD_{600} against water as a blank.
3. Transfer the overnight culture to a 150-ml shaker flask containing 50 ml of M9/Casamino Acids supplemented with appropriate antibiotic.
4. Take an initial 1-ml sample as a control for uninduced cells and either immediately treat as described in step 7 or refrigerate at –20°C until ready to perform step 7.
5. Incubate the culture for 1 to 2 hr in a 37°C orbital shaker with vigorous agitation (200 rpm). Check OD_{600}.

   The cells should be in full exponential growth phase (OD_{600} of 0.2 to 0.5).
6. Induce protein expression by adding IAA (2.5 ml of the 400× stock per liter) to a final concentration of 25 µg/ml. Incubate for another 3 hr and take a second (final) 1-ml sample.
7. For each 1-ml sample, microcentrifuge 2 min at maximum speed, room temperature. Discard supernatant and prepare the sample for SDS-PAGE analysis (see Support Protocol 4).

   If it is not convenient to electrophorese samples right away, pellets can be stored at –20°C.
8. Use remaining culture for analysis of protein solubility (see Support Protocol 5), and localization of proteins in the periplasmic space (see Support Protocol 6) or culture medium (see Support Protocol 7), or analysis of plasmid stability (if needed; see Support Protocol 2). If desired, store leftover cells (see Support Protocol 3).

ALTERNATE PROTOCOL 1

EXPRESSION UNDER CONTROL OF *lac*/tac PROMOTERS

For protein expression under control of the *lac* promoter or tac promoter (a hybrid between *lac* and *trp* promoters), induction via release of the *lac* repressor is achieved by adding isopropyl-β-thiogalactopyranoside (IPTG) to cultures, which binds and inactivates the repressor. The host strain must harbor the *lacI* repressor gene in the chromosome or in a plasmid (either the same plasmid as the target gene or a compatible one).

***Additional Materials** (also see Basic Protocol 2)*

Host *E. coli* strain containing *lacI* repressor gene and transformed with plasmid containing *lac* or tac promoter and gene for the protein of interest
Isopropyl-β-thiogalactopyranoside (IPTG; see recipe)

For this protocol, follow steps 1 to 5 of Basic Protocol 2 (*trp* promoter), substituting the appropriate host strain as the starting material. Replace step 6 with the following and then return to steps 7 and 8 of Basic Protocol 2.

6a. Induce protein expression by adding IPTG to a final concentration of 0.4 to 1.0 mM. Incubate for a further 3 hr and take a second (final) 1-ml sample.

BSU IASRI / CENTRAL UNIVERSITY LIBRARY

T7 RNA POLYMERASE/PROMOTER EXPRESSION SYSTEM

ALTERNATE PROTOCOL 2

The T7 RNA polymerase/promoter expression system utilized in this protocol involves induction, via the *lacUV5* promoter, of the highly active and efficient T7 RNA polymerase, which binds to the T7 promoter (p_{T7}) and transcribes genes under its control. An appropriate *E. coli* strain—in which the T7 polymerase gene under control of *lacUV5* has been integrated into host DNA and which has been transformed with plasmid DNA containing an ampicillin-resistance gene and the gene for the protein of interest under control of p_{T7}—is induced with IPTG. In a variation of the system, T7 lysozyme, which binds the polymerase and thus eliminates transcription under noninduced conditions, can be introduced on a second plasmid (e.g., pLysE or pLysS, which also encode chloramphenicol resistance) and maintained under selection with a second antibiotic.

Additional Materials (also see Basic Protocol 2)

Host *E. coli* strain transformed with DNA encoding T7 RNA polymerase under control of *lacUV5* promoter, plasmid containing p_{T7} and gene for the protein of interest, and (optional) plasmid containing gene for T7 lysozyme
Isopropyl-β-thiogalactopyranoside (IPTG; see recipe)

For this protocol, follow steps 1 to 5 of Basic Protocol 2, except begin with the appropriate transformed *E. coli* host strain, use ampicillin selection, and include a second antibiotic if the optional T7 lysozyme plasmid strategy is used (e.g., chloramphenicol at 30 mg/liter for pLysE or pLysL transformation). Replace step 6 with the following and return to steps 7 and 8 of Basic Protocol 2.

6b. Induce protein expression by adding IPTG to the culture to achieve a final concentration of 0.5 to 1.0 mM (10 to 20 ml of 100× stock per liter). Incubate for another 3 hr before taking a second (final) 1-ml sample.

TRANSFORMATION BY ELECTROPORATION

SUPPORT PROTOCOL 1

Electroporation is a quick, simple, and efficient method for introducing foreign DNA into *E. coli* (i.e., transformation). Cultures of the appropriate host strain are grown to exponential phase in rich medium to produce competent cells (i.e., cells able to take up foreign DNA). Competent cells can be stored frozen at –80°C until use. Plasmid DNA is introduced by subjecting cells to an electrical pulse, using a commercially available electroporation apparatus, and cells are allowed to recover in rich medium supplemented with glucose. After the transformation, strains can be stored (see Support Protocol 3) and/or used for protein expression (see Basic Protocol 1, Basic Protocol 2, Alternate Protocol 1, and Alternate Protocol 2). Plasmid stability analysis (see Support Protocol 2) is not absolutely required and should be carried out only in cases where expression with a given strain is not reproducible. Several different host strains can be transformed in the same experiment to allow investigation of the influence of the host on protein expression efficiency.

Materials

Host *E. coli* strain
LB medium (see recipe)
Water, ice-cold
10% (w/v) glycerol, ice-cold
Plasmid DNA
SOC medium (see recipe)
LB plates with appropriate antibiotic (see recipe)

BHARATHIAR UNIVERSITY LIBRARY

1-liter and 50-ml centrifuge bottles, chilled
Centrifuge (e.g., Sorvall RC3B equipped with H6000A rotor, or equivalent), 4°C
1.5-ml sterile microcentrifuge tubes, chilled
Electroporation apparatus (e.g., *E. coli* Pulser Apparatus; Bio-Rad)
Electroporation cuvettes, 0.2-cm electrode gap, sterile and prechilled

Prepare the competent cells

1. Prepare an overnight culture with the *E. coli* host strain to be transformed by inoculating a 100-ml flask containing 20 ml LB medium with a single colony from an LB agar plate. As the host strain is plasmid free, do not add any antibiotic; with shaking, grow at 37°C. The next morning, verify that cultures have reached stationary growth phase (OD_{600} of 1 to 2).

   In the case of T7 expression and when using the additional pLys plasmids (optional; see Alternate Protocol 2), the host already contains a plasmid and should be kept under selective pressure (30 mg/liter chloramphenicol).

2. Transfer the overnight culture to 800 ml LB medium in a 3-liter Erlenmeyer flask. Incubate cells 1 to 2 hr in a 37°C orbital shaker with vigorous agitation (200 rpm).

3. Verify that the cells are in full exponential growth phase (OD_{600} of 0.5 to 0.6). Transfer cells to a 1-liter chilled centrifuge bottle.

4. Centrifuge 15 min at 4000 × *g*, 4°C. Discard supernatant and resuspend the pellet in 5 to 10 ml ice-cold water. Bring to a final volume of 800 ml with ice-cold water.

5. Repeat step 4 twice.

6a. *If cells are to be transformed the same day:* Resuspend the pellet in ice-cold water to a final volume of 50 ml, and pellet again in a chilled 50-ml centrifuge bottle. Discard supernatant, estimate the volume of the pellet, and add an equal volume of ice-cold water. Vortex, distribute 100-µl aliquots into prechilled 1.5-ml sterile microcentrifuge tubes, and keep on ice. Proceed to step 7.

6b. *If competent cells are to be stored frozen:* Resuspend the pellet in ice-cold 10% glycerol to a final volume of 50 ml, and pellet again in a 50-ml centrifuge bottle. Discard supernatant, estimate the volume of the pellet, and add an equal volume of ice-cold 10% glycerol. Vortex, place 100-µl aliquots into prechilled 1.5-ml microcentrifuge tubes, and store in a −80°C freezer.

   To use frozen competent cells for transformation, thaw microcentrifuge tubes on ice and proceed to step 7.

Transform the competent cells

7. Add 5 pg to 0.5 µg plasmid DNA (in 1 to 5 µl) to 100 µl competent cells. Mix by inverting the tube several times by hand. Keep on ice.

   Quantities of plasmid DNA can be estimated from average yields of a plasmid prep, or measured by UV absorbance.

   Some host cells (such as E. coli B) require more DNA for efficient transformation, depending on their ability to restrict foreign DNA.

8. Set the electroporation apparatus to 2.5 kV, 25 µF, 200 Ω.

   Some variation of the resistance (100 to 400 Ω) may be necessary to obtain optimal results (see Seidman et al., 1989).

9. Transfer the DNA and cells to a prechilled electroporation cuvette, wipe away ice and

BSU IASI CENTRAL UNIVERSITY LIBRARY

water from the cuvette outer walls, place the cuvette into the electroporation chamber, and apply the electrical pulse.

Include a control with no DNA to ensure that transformants arise only by introduction of foreign DNA into the competent cells.

10. Remove the cuvette, immediately add 1 ml SOC medium and transfer the contents of the cuvette to a 20-ml sterile culture tube. Incubate 60 min with moderate shaking.

    The incubation temperature depends on the expression vector. Use 30°C for p_L-based expression plasmids and 37°C for others.

11. Spread aliquots of the transformation culture (10 to 100 µl) on LB plates with appropriate antibiotic. Check for appearance of individual colonies after 24 to 48 hr of incubation.

12. Individually streak and label 5 to 10 colonies from transformation plates onto fresh plates with appropriate antibiotic to ensure they are pure, viable, and antibiotic resistant.

13. Store an aliquot of each strain (see Support Protocol 3). Proceed to the appropriate expression protocol (see Basic Protocol 1, Basic Protocol 2, Alternate Protocol 1, and Alternate Protocol 2).

PLASMID STABILITY ANALYSIS

SUPPORT PROTOCOL 2

Analysis of plasmid stability is important for assessing the quality of recombinant strains. This check is not absolutely required and is really only needed in cases where expression with a given strain is not reproducible. An aliquot of the frozen stock is thawed and serial dilutions of the strain are prepared and spread on plates containing rich medium with or without the antibiotic used to maintain the plasmid. Plated cultures are grown overnight, colonies counted, and the percentage of plasmid-carrying cells in the population determined.

Materials

PBS (*APPENDIX 2E*)
Transformed *E. coli* host strain (Basic Protocols 1 and 2, Alternate Protocols 1 and 2)
LB plates with and without antibiotic (see recipe)
20-ml sterile culture tubes
Sterile triangular glass rod

NOTE: Cells should be kept in an ice bath throughout the procedure to minimize alterations to the samples.

1. Set up a rack of 20-ml sterile culture tubes and place 9 ml PBS in each tube. Pipet a 1-ml sample of the transformed *E. coli* host strain to be analyzed into the first tube. Vortex for 5 sec.

   An OD_{600} reading should be done before diluting the culture to facilitate steps 2-3.

2. With a new sterile pipet, transfer 1 ml from the previous tube into the next tube containing PBS.

3. Repeat step 2 until the final desired dilution is reached.

   The dilution range depends on the initial density of the culture. For an OD_{600} of 1, the final dilution factor should be 10^{-6}.

BCU LAS / CENTRAL UNIVERSITY LIBRARY

4. Spread 100 µl of the final three dilutions onto triplicate sets of LB plates with and without antibiotic. Use a sterile triangular glass rod to achieve an even distribution of the cells on the plate.

   Start with the highest dilution (i.e., with the fewest cells) to minimize effects of carryover from one plate to the next. As 0.1 ml is spread on the plate and as a culture at an OD_{600} of 1 contains ~10^9 cells/ml, diluting the culture to a maximum of 10^{-6} should give ~100 colonies per plate.

5. Incubate plates overnight at the appropriate temperature.

   A temperature of 30°C should be used for p_L-based expression strains and 37°C for all others. Incubate the plates upside down to collect water condensation in the plate cover.

6. The next morning, count the colonies on each plate. Calculate the number of cells present in the original sample that grew on LB and the number that grew on LB plus antibiotic. Determine the ratio of the two numbers to obtain the percentage of plasmid-carrying cells in the population.

   Serial dilutions can be made by a different combination of volumes leading eventually to a different dilution factor (0.1 ml into tubes containing 0.9 ml, 0.01 ml in 0.99 ml, etc.) For best accuracy, count plates with >30 colonies.

   Strains showing more than 50% plasmid retention before induction of protein expression are acceptable.

SUPPORT PROTOCOL 3

STRAIN STORAGE

For long-term storage, recombinant *E. coli* strains are kept at low temperature, frozen with a cryopreservative agent. A temperature of –80°C is sufficient for most strains. Multiple strains can be prepared in parallel and frozen in a single operation. It is good practice to keep at least one frozen copy that is not thawed and frozen repeatedly. The frozen stocks can be retrieved from storage and used for protein expression (Basic Protocols 1 and 2, and Alternate Protocols 1 and 2).

Materials

Transformed *E. coli* host strain (see Support Protocol 1)
LB medium (see recipe)
Antibiotic stock solution (see recipe)
20% (w/v) glycerol, dispersed in sterile cryovials (see recipe)
LB plate (see recipe)

20-ml culture tube
30° or 37°C orbital shaker

1. Grow an overnight culture of transformed *E. coli* host strain by inoculating 2 ml of LB medium (supplemented with appropriate antibiotic, if necessary) in a 20-ml culture tube with a single colony from a fresh agar plate. Incubate the tube overnight on an orbital shaker. The next morning, verify that cells are in stationary phase (OD_{600} of 1 to 2).

   The incubation temperature is dependent on the expression vector. Use 30°C for p_L-based systems and 37°C for others.

2. Inoculate a 20-ml tube (containing 2 ml LB plus appropriate antibiotic) with 50 µl from the overnight culture. Incubate for 2 to 3 hr.

ESU IASI / CENTRAL UNIVERSITY LIBRARY

3. Add 0.9 ml culture to each cryovial, label, vortex well, and freeze immediately. Store frozen strains at –80°C.
4. To start a new culture from the frozen stock, scrape the surface of the vial with a sterile loop and either streak on an LB plate or dip into liquid medium. Return the cryovial to –80°C for storage.

Several hundred strains can be stored in a single freezer, but it is advisable to have alarm and safety devices against mechanical or electrical failure.

PREPARATION OF SAMPLES FOR ANALYSIS BY SDS-PAGE

SUPPORT PROTOCOL 4

Electrophoretic separation of proteins in a cell extract (total lysate or fraction of the cell) by SDS-PAGE is described in detail in *UNIT 10.1*. This protocol describes only the preparation of the samples and applies to material derived from Basic Protocol 1, Basic Protocol 2, Alternate Protocol 1, and Alternate Protocol 2. The objective is to bring all proteins of the sample into solution in a denatured state so the mixture can be resolved by one-dimensional electrophoresis in SDS gels. Solubilization is achieved by sonicating samples in a buffer containing detergent and reducing agents, followed by brief heating. A large excess of sample is prepared so that it can be reloaded on several gels, if necessary. Once solubilized in SDS sample buffer, samples can be stored for years at –20°C.

Materials

Sample of interest
2× and 1× SDS sample buffer (*UNIT 10.1*)

Toothpicks, sterile
Sonicator (Branson, with microtip)

1a. *For soluble samples (e.g., soluble cellular proteins, periplasmic extracts, extracellular medium):* Add the sample of interest to an equal volume of 2× SDS sample buffer in a 1.5-ml microcentrifuge tube and proceed directly to step 4.

1b. *For pellet samples (total cells, insoluble fraction, trichloroacetic acid precipitate):* Transfer 5 to 10 mg pellet material from the sample of interest with a sterile toothpick to a preweighed 1.5-ml microcentrifuge tube.

2. Add 1× SDS sample buffer, using 400 µl per 10 mg pellet.
3. Sonicate 10 sec with a sonicator (with power setting at 30 to 50 W) to resuspend the pellet.

Be careful to avoid foaming by keeping the sonicator probe well below the liquid surface at all times.

4. Heat to 90°C for 5 min in a heating block.

After cooling, samples are ready to be applied to gel wells for electrophoresis (UNIT 10.1). Loading amounts are based on gel and well dimensions, as well as the detection system to be used (e.g., silver staining is 100 times more sensitive than Coomassie blue staining; UNIT 10.5). Recommended amounts are as follows: 0.5 to 1 µl for microgels (e.g., 4.3 × 5 cm; 0.45 mm thick), 5 to 10 µl for minigels (e.g., 7.3 × 8.3 cm; 1 mm thick), and up to 20 µl for full-size gels (e.g., 14 × 14 cm; 1.5 mm thick).

B. C. U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

SUPPORT PROTOCOL 5

SOLUBILITY ANALYSIS

Analysis of the solubility of recombinant proteins is needed to determine the subsequent purification strategy (*UNIT 6.1*). The protocol is based on mechanical disruption of cells followed by high-speed centrifugation to separate the soluble fraction from the insoluble fraction. Samples are then prepared for SDS-PAGE analysis (see Support Protocol 4).

Materials

Transformed *E. coli* host cells expressing the protein of interest
25 mM HEPES (pH 7.6; autoclave and store up to 1 yr at 4°C)
50- and 5-ml centrifuge tubes
Centrifuge (e.g., Sorvall RC5B equipped with SS34 rotor), 4°C
Sonicator

NOTE: Cells should be kept in an ice bath throughout the procedure to minimize alterations to the samples.

1. Harvest 25 ml transformed *E. coli* host cells expressing the protein of interest by centrifuging 15 min at 4000 × *g*, 4°C, in a 50-ml centrifuge tube.
2. Discard the supernatant and resuspend the pellet in 1 to 1.5 ml 25 mM HEPES (pH 7.6) to obtain an OD_{600} of ~40.
3. Set the tube in an ice bath and sonicate three times for 1 min each, allowing 30-sec intervals between pulses.
4. Place a 100-µl aliquot of the cell lysate (total protein) in a 1.5-ml microcentrifuge tube and keep at –20°C for later analysis (step 7).

   The 30-sec intervals allow the suspension to cool down between sonications. Be careful to avoid overheating at all stages.

5. Transfer the suspension to a 5-ml centrifuge tube and centrifuge 30 min at 30,000 × *g* (16,000 rpm in an SS-34 rotor), 4°C.
6. Pour off supernatant (soluble fraction) into another 5-ml tube to separate it from the pellet (insoluble fraction).
7. Prepare the three fractions (total, soluble, and insoluble proteins) for analysis by SDS-PAGE (see Support Protocol 4).

   The proportion of soluble/insoluble proteins can be quantified after SDS-PAGE separation (UNIT 10.1), staining of gels (UNIT 10.5), and scanning densitometry of the gels.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

PREPARATION OF PERIPLASMIC EXTRACTS

This protocol describes a small-scale procedure to confirm the expression of proteins that are designed to be localized to the periplasmic space of *E. coli*. It is based on subjecting cells to an osmotic shock in a concentrated sucrose solution, which mechanically forces the contents of the periplasm through the outer membrane and cell wall.

Materials

Transformed *E. coli* cells expressing the protein of interest
PBS (*APPENDIX 2E*)
20% (w/v) sucrose/10 mM Tris·Cl (pH 7.5), ice-cold
0.5 M EDTA (pH 8.0)
Ice-cold water

Microcentrifuge, 4°C

NOTE: Cells should be kept in an ice bath throughout the procedure to minimize alterations to the samples.

Release the periplasmic proteins by osmotic shock

1. Place 1 ml transformed *E. coli* cells expressing the protein of interest in a 1.5-ml microcentrifuge tube and microcentrifuge 2 min at maximum speed, 4°C.

   Depending on the culture density, adjust the volume of the sample so that it contains ~2 × 10^9 cells. (An OD_{600} of 1 is equivalent to 10^9 cells/ml).

2. Discard supernatant and resuspend the pellet in 1 ml PBS. Vortex tube well.

3. Pellet the cells again in the microcentrifuge for 2 min.

4. Discard supernatant and resuspend the pellet in 150 µl cold 20% sucrose/10 mM Tris·Cl (pH 7.5). Vortex well and add 5 µl of 0.5 M EDTA (pH 8.0).

5. Save a 50-µl aliquot representing the whole-cell fraction and store at –20°C. Leave remainder of the sample on ice for 10 min.

   The short incubation allows time for the EDTA to loosen the cell wall structure.

Separate the sample into periplasmic and cytoplasmic/membrane fractions

6. Pellet the cells for 5 min in the microcentrifuge.

7. Discard supernatant and resuspend the pellet by vigorous vortexing in 100 µl ice-cold water.

8. Leave on ice for 10 min and then repeat step 6.

9. Carefully collect and save the supernatant, which contains the periplasmic fraction in a 1.5-ml microcentrifuge tube and store at –20°C.

10. Eliminate remaining fluid from the pellet with a cotton-tipped swab and resuspend the sample in 100 µl ice-cold water. Save this solution as the cytoplasmic and membrane fraction.

11. Prepare the three fractions (whole-cell, periplasm, and cytoplasm/membrane fractions) for SDS-PAGE analysis (see Support Protocol 4).

BCU Main Central University Library

PREPARATION OF EXTRACELLULAR MEDIUM SAMPLES

Expression systems that direct extracellular secretion of the desired protein can be checked by analyzing samples of the culture medium. This protocol describes treatment of culture media with trichloroacetic acid (TCA) to precipitate extracellular proteins and thus concentrate proteins for subsequent analysis. Addition of bovine serum albumin (BSA) may be helpful if protein concentrations in the medium are low.

Materials

Transformed *E. coli* cells expressing the protein of interest
2 g/liter BSA (optional; sterilize through a 0.22-µm filter and store ≤1 month at 4°C)
200 g/liter trichloroacetic acid (TCA; sterilize through a 0.22-µm filter and store ≤1 month at 4°C)
70% (v/v) ethanol
0.22-µm filters suitable for aqueous solutions
Microcentrifuge, 4°C
Ice bath

1. Place a 1-ml culture of *E. coli* cells expressing the protein of interest in a 1.5-ml microcentrifuge tube and microcentrifuge 2 min at maximum speed, 4°C.
2. Collect the supernatant and pass it through a 0.22-µm filter into a 1.5-ml microcentrifuge tube. Discard pellet.
3. Add 100 µl of 2 g/liter BSA, if necessary.

   BSA acts as a carrier and nucleus for precipitation. It can be omitted if the medium already contains >100 mg/l of protein.
4. Add 1 ml of 200 g/liter TCA, vortex well, and leave the tube on ice for 15 to 30 min to allow proteins to precipitate.
5. Microcentrifuge for 10 min at maximum speed, 4°C.
6. Discard supernatant and wash the pellet at least twice with 2 ml of 70% ethanol.

   The washings remove the TCA, which would otherwise disturb subsequent analysis by SDS-PAGE.
7. Prepare the pellet for analysis by SDS-PAGE (see Support Protocol 4).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Antibiotic stock solutions (1000× concentrated)

Ampicillin: Prepare a 40 g/liter stock in water. Store ≤2 weeks at 4°C.

Chloramphenicol: Prepare a 30 g/liter stock in ethanol. Store ≤1 year at –20°C.

Kanamycin: Prepare a 40 g/liter stock in water. Store ≤1 year at 4°C.

Tetracycline: Prepare a 5 g/liter stock in 70% ethanol. Store ≤1 year in the dark at –20°C.

Sterilize all stock solutions by filtration through a 0.22-µm filter before use. Aliquot in 5 to 10 portions, depending on quantities needed.

BSU IASRI CENTRAL UNIVERSITY LIBRARY

Glycerol

Dispense 1 ml of 20% (v/v) glycerol into 2.5-ml cryovials. Autoclave with the caps loose. After sterilization, tighten the caps. Store up to 1 year at 4°C.

An entire rack of cryovials can be prepared at once, as they can be stored for a long time.

Indole-3-acrylic acid (IAA)

Store IAA as a powder at 4°C. Prepare solution fresh as a 400× concentrated solution in ethanol (10 g/liter).

Isopropyl-β-thiogalactopyranoside (IPTG)

Store IPTG as a powder at −20°C. Prepare solution fresh as a 100× concentrated solution in water (50 mM, i.e., 11.9 mg/liter). Sterilize by filtration before use.

LB medium

Per liter:
10 g Bacto-tryptone
5 g yeast extract
5 g NaCl
10 ml 1 N NaOH (to pH 7.0)
H_2O to 1 liter
Sterilize 30 min at 121°C

Dispense into 20-ml culture tubes (2 or 5 ml/tube) or shaker flasks (see Commentary). Plug opening with rigid aluminum cap or with hydrophobic cotton wool covered with aluminum foil.

Antibiotics should be added (if appropriate) just before use.

The original recipe for LB medium (sometimes referred to Luria or Lenox broth) does not contain NaOH. There are many different recipes for LB that differ only in the amount of NaOH added. We use this formula in our own work. Even though the pH is adjusted to near 7.0 with NaOH, the medium is not highly buffered, and the pH of a culture growing in it drops as it nears saturation.

M9 minimal medium

Per liter:
6 g Na_2HPO_4
3 g KH_2PO_4
1 g NH_4Cl
0.5 g NaCl
3 mg $CaCl_2$
H_2O to 1 liter
Sterilize 30 min at 121°C

Dispense into 20-ml culture tubes (2 or 5 ml/tube) or shaker flasks (see Commentary). Plug opening with rigid aluminium cap or with hydrophobic cotton wool covered with aluminum foil.

Antibiotics should be added (if appropriate) just before use.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

SOC medium

Per liter:

5 g yeast extract
20 g Bacto-tryptone
580 mg NaCl (10 mM)
186 mg KCl (2.5 mM)
940 mg $MgCl_2$ (10 mM)
1.2 g $MgSO_4$ (10 mM)
3.6 g glucose (20 mM)
H_2O to 1 liter
Sterilize 30 min at 121°C

Dispense into 20-ml culture tubes (2 or 5 ml/tube) or shaker flasks (see Commentary). Plug opening with rigid aluminium cap or with hydrophobic cotton wool covered with aluminum foil.

Antibiotics should be added (if appropriate) just before use.

The amounts given in the recipe refer to the anhydrous salts.

Plates

Use liquid medium formulation (see recipes for LB, SOC, and M9 media) and add 20 g/liter agar.

For LB or SOC plates, autoclave the agar together with the other ingredients. For M9 plates, autoclave 20 g agar in 800 ml water for 15 min, then add sterile concentrated M9 medium (i.e., make up components in 200 rather than 1000 ml water). Cool the agar solutions to about 50°C before adding desired supplements (i.e., antibiotics; see recipe). The media will stay liquid at 50°C but will rapidly solidify if the temperature drops below 45°C. Pour 32 to 40 ml medium into each sterile disposable petri dish to get 25 to 30 plates per liter. Allow the agar to solidify. Dry the plates by leaving them out at room temperature for 2 or 3 days, or by leaving them with the lids off for 30 min in a 37°C incubator or in a laminar-flow hood. Store dry plates at 4°C, wrapped in the original bags used to package the empty plates. Antibiotic-containing plates have a limited shelf-life: 2 weeks for ampicillin; 6 weeks for chloramphenicol, kanamycin, and tetracycline.

COMMENTARY

Background Information

As described in *UNIT 5.1*, the choice of an expression system and corresponding host strain for the production of a recombinant protein in *E. coli* is dictated by the ultimate goal of the study. If the protein is needed for structural studies requiring tens to hundreds of milligrams of a given protein, correctly folded, then either secretion of the protein or high-level production inside the cell (followed by protein refolding if the protein is in aggregated form; *UNIT 6.5*) may be the methods of choice. If an antigen is needed for raising antibodies, gene fusion strategies are often the method of choice. The toxicity of some recombinant proteins will also drive a requirement for very tightly repressed systems such as T7 RNA polymerase, which have particular host requirements (see Alternate Protocol 2). In other words, the type of protein and its solubility dictate the vector construction and, thus, the *E. coli* host strain used.

The p_L promoter from lambda phage initiates high levels of transcription that can be repressed by the cI protein (Shatzman et al., 1990). A temperature-sensitive mutant of the repressor, cI857, which is inactive at high temperature, allows induction by a temperature shift from 30° to 42°C (Remaut et al., 1981). The repressor gene can be carried on the host chromosome or on the same plasmid as p_L. In general, induction of p_L is most efficient when the cells are in full exponential growth phase in a rich medium (see Basic Protocol 1).

Tac is a hybrid between *lac* and *trp* promoters and is regulated by the same *lacI* repressor as *lac*. These promoters are inducible by the addition of a nonmetabolizable galactose analog, isopropyl-β-thiogalactopyranoside

ECU IASI - CENTRAL UNIVERSITY LIBRARY

(IPTG), which binds and inactivates the repressor. The *lacI* repressor gene can be integrated in the host chromosome, on the same plasmid as the target gene, or on another compatible one. This expression system (see Alternate Protocol 1) is extensively used for recombinant protein expression (Harris, 1983).

T7-driven expression (see Alternate Protocol 2) is based on the efficiency and specificity of the phage T7 RNA polymerase for initiating transcription from T7 promoters (Tabor, 1990; Studier and Moffat, 1986). Two elements are introduced into the host cell: the T7 polymerase gene, under control of an inducible promoter (*lacUV5*), which is integrated into the host DNA as a λDE3 lysogen; and the target gene, which is introduced on a plasmid under control of the T7 promoter. This plasmid also carries an antibiotic (ampicillin or kanamycin) resistance gene for selection. Two hosts are routinely used: *E. coli* BL21 and HMS174 (Table 5.2.1). For toxic target gene products and because basal levels of the T7 polymerase can promote some transcription of the target gene, a third optional element—T7 lysozyme—is sometimes added. T7 lysozyme binds the T7 polymerase, thereby inactivating it and eliminating transcription of the target gene under noninduced conditions. The T7 lysozyme gene is carried by either of two plasmids (pLysE or pLysS) and expressed constitutively. Both plasmids also carry a selection marker (chloramphenicol-resistance gene). The pET system is a series of vectors for cloning into the T7 promoter system (Novagen). Some vectors carry an ampicillin-resistance gene; others carry kanamycin resistance. Novagen distributes a complete set of tools associated with the system, including host strains, competent cells, and pLysS and pLysE vectors.

The promoter for the tryptophan operon (*trp*) is strong and well-repressed under basal conditions (Somerville, 1988). Tryptophan present in the culture medium acts as a corepressor by binding to the *trp* repressor. This complex binds to the operator and prevents transcription. As the cells consume tryptophan for growth, its concentration decreases to a point where the operator is unbound and transcription of genes downstream (see Basic Protocol 2) is initiated. Addition of indole-3-acrylic acid (IAA), a tryptophan structural analog, inactivates any residual repressor.

High-copy-number plasmids are desirable for high-level expression but also often result in plasmid loss. An analysis of plasmid stability (see Support Protocol 2) allows evaluation of plasmid loss and control of the quality and integrity of recombinant strains.

The relative proportions of soluble and insoluble recombinant protein (also referred to as inclusion bodies) are affected by many variables such as protein primary sequence, expression system, and host strain. Environmental factors such as temperature and pH of the culture are also key determinants as there is a consensus that lower temperatures (20° to 25°C) lead to a higher proportion of soluble protein. The soluble/insoluble ratio dictates the subsequent purification strategy (see Chapter 6). It is therefore essential to analyze the recombinant protein solubility (see Support Protocol 5), which, in some cases, can be considered more important than the total protein expression level. Secretion to the periplasmic space should be evaluated, because this compartment provides a more oxidative environment than the cytoplasm and favors disulfide bridge formation.

With an appropriate expression system, it is possible to direct secretion of the recombinant protein to the periplasmic space of *E. coli* (Hsiung et al., 1986). Analysis of expression and confirmation of the proper localization thus require the preparation of a periplasmic extract. The osmotic shock method (see Support Protocol 6) to prepare periplasmic extracts was originally described by Koshland (1980). Fusion to periplasmic (e.g., MalE, PhoA) or outer membrane (e.g., OmpA, LamB) protein signal peptides has been extensively used to direct recombinant proteins to the periplasm (Hsuing et al., 1986; Hewinson and Russell, 1993; Karim et al., 1993). This approach was successfully applied to antibody fragments (Carter et al., 1992; Plückthun, 1991). Complete secretion to the extracellular medium has also been achieved by fusing the gene of interest at the N terminus to protein A signal peptide from *S. aureus* or at the C terminal domain of hemolysin A (reviewed in Blight and Holland, 1994).

Several expression systems are now available that allow efficient secretion of the recombinant protein into the extracellular medium (Blight and Holland, 1994). However, the secreted protein is usually quite dilute and concentration of the protein is necessary for analyzing and confirming expression by SDS-PAGE. Ultrafiltration or precipitation can be used for this purpose. This protocol describes precipitation by TCA, which is quick, reliable, and inexpensive. Several samples can be handled in the same batch.

PUBLISHED BY THE CENTRAL LIBRARY

Critical Parameters and Troubleshooting

The temperature shift necessary to induce p_L will also induce a heat-shock response, and in particular turn on some protease genes. Also, some host strains (such as *E. coli* HB101) do not tolerate 42°C, and growth is drastically reduced at this temperature. Tac and *lac* are weaker than p_L or *trp*, but still drive high-level expression of some proteins. They are not recommended for toxic gene products because repression is often incomplete. This occurs because, in the *lac* operon, normal repression is mediated in part by an operator site internal to the *lacZ* gene, which is not usually found in the target gene. It is also important to note that repression is more efficient with the mutated repressor gene $lacI^Q$.

The expression protocols can be scaled to any size shaker flask, as long as the culture is kept well aerated (maximum ratio of liquid/flask volume of 0.3). Care should also be taken to ensure that incubation temperatures are accurate. It takes ~95 min for an 800-ml culture in a 2-liter flask to heat from 20° to 42°C in an air shaker at 200 rpm. Pre-warming of culture media can be done adequately in an air shaker if enough time is allowed for the warmup, but it is probably better if a water bath is used for temperature induction steps.

Attempts to increase solubility of proteins formed in aggregates or inclusion bodies have included coexpression with chaperones, manipulation of temperature and other environmental factors, and secretion of the desired protein to the periplasmic space and/or the extracellular medium. These strategies have been reviewed (Hockney, 1994).

Incomplete repression is quite frequent with genes under control of a *lac*-type promoter. In fact, for some cases, one may find that induction does not increase the constitutive level of expression. This is due in part to the mechanism of repression by the LacI protein. It has been shown that the repressor binds the natural *lac* operator in a DNA region that also contains part of the 5′ coding sequence for LacZ. If this region is exchanged for that of a foreign gene, then the affinity of the repressor for the operator is decreased and transcription can occur at a reduced level.

The optional pLysE or pLysS plasmids for the T7 system should be used with considerable caution. The presence of pLysE leads to a high level of lysozyme in the cell, which can dampen expression and eventually weaken the cell wall. It also, however, brings an additional advantage for purification: a simple freeze-thaw cycle usually effects cell lysis efficiently.

In the case of no detectable expression, we advise (1) trying another promoter system; (2) testing several different host strains; (3) fusing the gene of interest to well-expressed DNA sequences such as LacZ, glutathione transferase (GST), thioredoxin, or protein A, to aid expression and/or recovery (*UNIT 5.1*).

Optimization of codon usage is a topic outside the scope of this unit (Robinson et al., 1984). Overexpression of rare tRNAs and mutagenesis of rare codons have been used for enhancing expression of heterologous proteins in *E. coli* (De Boer and Hui, 1990).

DNA used for electroporation should be as free of salts as possible, as they contribute to decreasing the resistance of the sample and can lead to arcing. It is quite useful to make large batches of competent cells (e.g., 800-ml cultures) and freeze them in aliquots. Several host strains can be transformed quickly (in 2 hr) tested in parallel for expression. It should also be stressed that the goal of the transformation is only to obtain a few colonies of *E. coli* cells carrying the expression vector. Efficiency is not a major concern. Usually a few hundred transformant colonies are generated; only a small proportion is tested for the ability to express the recombinant protein of interest.

SDS-PAGE analysis of total *E. coli* cell proteins is straightforward and reproducible. One minor difficulty may arise from pellet sample preparation—if DNA has not been broken down well enough at the sonication step, the sample will be highly viscous and hard to load on the gel. In this instance, sonication should be repeated. At low levels of expression, it is possible that Coomassie staining will not be sensitive enough. It is not recommended that more sensitive staining methods (e.g., silver staining) be tried, as this will only increase the background. Rather, immunoblotting should be used to analyze the proteins. It is also good practice to include negative controls, such as noninduced cell extract or extract of induced cells carrying an empty vector, as well as positive controls, such as samples from previous experiments or partially pure protein where available.

Anticipated Results

Expression levels of the desired proteins may range from <1% to >40% of total cell protein. In shaker-flask cultures this is equivalent to hundreds of milligrams per liter, which can be increased to grams-per-liter level in

BSU IASRI CENTRAL UNIVERSITY LIBRARY

fermenters (where greater cell densities are achieved; see *UNIT 5.3*). The expressed product can usually be visualized by running appropriate extracts of cells on SDS-polyacrylamide gels and staining with Coomassie blue. The majority of the proteins produced will be insoluble, but the material can often be solubilized and renatured to an active state (*UNITS 6.3, 6.4, & 6.5*).

Time Considerations

Transformation of cells requires about 1 day. Cell growth, induction, and harvesting take 6 to 8 hr depending on the expression system and strain of *E. coli.* Analyzing plasmid stability if necessary extends into the following day. It is often convenient to conduct analysis of samples by SDS-PAGE the next day, as gel preparation, running, staining, and destaining will require several hours. The requirement to prepare overnight cultures to be induced or analyzed must also be taken into account. Gel preparation can be avoided by using commercially available gels (e.g., Phast-gels from Pharmacia Biotech or minigels from Novex). Overall, 2 to 3 days should be required for full testing of a given expression vector in different host strains.

Literature Cited

Blight, M.A. and Holland, B. 1994. Heterologous protein secretion and the versatile *E. coli* haemolysin translocator. *Trends Biotechnol.* 12:450-455.

Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.-L.T., Rowland, A.M., Carver, M.E., Yang, M., Bourell, J.H., Shephard, H.M., and Henner, D. 1992. High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. *Bio/Technology* 10:163-167.

De Boer, H.A. and Hui, A.S. 1990. Sequences within ribosome binding site affecting messenger RNA translatability and method to direct ribosomes to single messenger RNA species. *Methods Enzymol.* 185:103-114.

Goeddel, D.V. (ed.) 1990. Gene expression technology. *Methods Enzymol.* 185.

Harris, T.J.R. 1983. Expression of eukaryotic genes in *E. coli. Genet. Eng.* 4:127-185.

Hewinson, R.G. and Russell, W.P. 1993. Processing and secretion by *Escherichia coli* of a recombinant form of the immunogenic protein MPB70 of Mycobacterium bovis. *J. Gen. Microbiol.* 139:1253-1259.

Hockney, R.C. 1994. Recent developments in heterologous protein production in *E. coli. Trends Biotechnol.* 12:456-463.

Hsiung, H.M., Mayne, N.G., and Becker, G.W. 1986. High level expression, efficient secretion and folding of human growth hormone in *E. coli. Bio/Technology* 4:991-995.

Karim, A., Kaderbhai, N., Evans, A., Harding, V., and Kaderbhai, M.A. 1993. Efficient bacterial export of a eukaryotic cytoplasmid cytochrome. *Bio/Technology* 11:612-618.

Koshland, D. and Botstein, D. 1980. Secretion of beta-lactamase requires the carboxy end of the protein. *Cell* 20:749-760.

Plückthun, A. 1991. Antibody engineering: Advances from the use of *Escherichia coli* expression systems. *Bio/Technology* 9:545-551.

Raleigh, E.A., Lech, K., and Brent, R. 1989. Selected topics from classical bacterial genetics. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 1.4.1-1.4.14. John Wiley & Sons, New York.

Remaut, E., Stanssens, P., and Fiers, W. 1981. Plasmid vectors for high-efficiency expression controlled by the p_L promoter of coliphage lambda. *Gene* 15:81-93.

Remaut, E., Marmenout, A., Simons, G., and Fiers, W. 1987. Expression of heterologous unfused proteins in *Escherichia coli. Methods Enzymol.* 153:416-431.

Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarranton, G., Stephens, P., Millican, A., Eaton, M., and Humphreys, G. 1984. Codon usage can affect efficiency of translation of genes in *Escherichia coli. Nucl. Acids Res.* 12:6663-6671.

Rosenberg, S.A., Grimm, E.A., McCrogan, M., Doyle, M., Kawasaki, E., and Koths, K. 1984. Biological activity of recombinant human interleukin-2 produced in *E. coli. Science* 223:1412-1415.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *In* Molecular Cloning, A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Seidman, C.E., Struhl, K., and Sheen, J. 1989. Introduction of plasmid DNA into cells. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 1.8.1-1.8.8. John Wiley & Sons, New York.

Shatzman, A.R., Gross, M.S., and Rosenberg, M. 1990. Expression using vectors with phage λ regulatory sequences. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.3.1-16.3.11. John Wiley & Sons, New York.

Somerville, R.L. 1988. The *trp* promoter of *E. coli* and its exploitation in the design of efficient protein production systems. *Biotechnol. Genet. Eng. Rev.* 6:1-41.

Stahl, S.J. and Murray, K., 1989. Immunogenicity of peptide fusions to hepatitis B virus core antigen. *Proc. Natl. Acad. Sci. U.S.A.* 86:6283-6287.

BIBLIOTECA CENTRAL UNIVERSITARIA LIBRARY

Studier, F.W. and Moffat, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *J. Mol. Biol.* 189:113-130.

Tabor, S. 1990. Expression using the T7 RNA polymerase/promoter system. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.2.1-16.2.11. John Wiley & Sons, New York.

Key References

Goeddel, 1990. See above.

Reviews most common expression systems in E. coli and provides a complete source of background information on the expression vectors as well as some hints on strategies for optimization.

Hockney, 1994. See above.

Comprehensive review on advances in protein expression in E. coli. Provides useful information regarding solubility, secretion, and fusions. Also discusses the difficult problem of expressing complex proteins such as multi-transmembrane proteins in E. coli.

Contributed by Alain Bernard and Mark Payton
Glaxo Institute for Molecular Biology
Geneva, Switzerland

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Fermentation and Growth of *Escherichia coli* for Optimal Protein Production

Large-scale production of recombinant proteins in *Escherichia coli* requires growth of cells in fermentors. Fermentation involves the transformation of simple organic compounds by the enzymatic machinery of *E. coli*, in this case to produce proteins. There are many choices of bacterial strain, vector, expression system, and protein localization and purification strategy that must be set before proceeding to batch fermentation. *UNIT 5.1* gives an overview of protein expression in *E. coli*, and *UNIT 5.2* outlines procedures for transforming cells and inducing synthesis of foreign proteins in various expression systems on a small scale in shaker flasks. Table 5.2.1 lists *E. coli* strains appropriate for use in fermentors. The Commentary of this unit discusses important characteristics of fermentation equipment.

Starting with a strain of *E. coli* that has been transformed to produce the desired protein (*UNIT 5.2*) this unit focuses on production of recombinant proteins in batch fermentations (Basic Protocol 1). Alternate protocols introduce variations of fermentation systems that enable continuous growth and protein production in high-cell-density, fed-batch cultures (Alternate Protocol 1) and that permit labeling of recombinant proteins with heavy atom derivatives such as selenomethionine (Alternate Protocol 2) or with stable isotopes such as ^{2}H, ^{13}C, and ^{15}N (Alternate Protocol 3). Production of labeled proteins facilitates structural studies by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. The protocols in this unit are designed for expression systems directing intracellular or periplasmic localization of recombinant proteins; however, in the case of extracellular secretion of the desired protein, the culture medium itself, rather than pelleted cells, would be saved, concentrated, and subjected to purification processes.

Fermentation experiments require careful monitoring of cell growth and assurance of preinoculation sterility, which are described in Support Protocols 1 and 2, respectively. It must be kept in mind, however, that the outcome of fermentation experiments, even if derived from successful small-scale shaker flask cultures (*UNIT 5.2*), may be negatively affected by unexpected toxicity of the target protein in fermentation systems.

PRODUCTION OF RECOMBINANT PROTEINS IN BATCH FERMENTATIONS

BASIC PROTOCOL 1

This protocol describes the basic sequence of operations, their timing, and the key issues associated with batch fermentations. Simple batch fermentations involve growth of fermenting organisms in a fixed amount of medium, without additional nutrients. The fermentor is set up, filled with fermentation medium (but not including heat-sensitive ingredients), and sterilized. Filter-sterilized heat-sensitive ingredients are then added, and environmental parameters are set to the desired values (using acid and base to adjust pH) before final preinoculation samples are taken to check sterility (Support Protocol 2). An inoculum prepared from a fresh overnight culture of the transformed *E. coli* strain is introduced into the fermentor and the cells allowed to grow. Synthesis of the recombinant protein following induction of cells is monitored, and cells are harvested by centrifugation using a batch pot or continuous feed centrifuge, depending on the volume of medium processed.

Special fermentation equipment is required. Several commercial manufacturers produce fermentation equipment (Table 5.3.1) for laboratory use (1 to 20 liters) or small pilot-scale experiments (50 to 200 liters). For recombinant *E. coli* processes, users should pay particular attention to heat and oxygen transfer, process monitoring and control (particu-

Copyright © 1995 by John Wiley & Sons, Inc.

EDINBURGH CENTRAL UNIVERSITY LIBRARY

Table 5.3.1 Fermentation Equipment Suppliers

Supplier[a]	Vessel size (liters)	Control	Comments
B. Braun	2 to 50,000	DDC[b], Computer supervision	Microprocessor-based instrumentation
Bioengineering	2.4 to 50000	Analog, DDC, Computer supervision	Special vessel designs (airlifts, fluidized beds, membrane dialysis) Complete plants
Inceltech	2.5 to 800	Analog, DDC, Computer supervision	Very high oxygen transfer
New Brunswick Scientific Co.	1 to 15,000	Analog, DDC, Computer supervision	Very high oxygen transfer Windows-based control package
LSL Biolafitte	1 to 50,000	DDC, Computer supervision	Microprocessor-based instrumentation
New MBR	3.5 to 100,000	Analog, DDC, Computer supervision	Complete plants

[a]For complete listing of suppliers' addresses, see *SUPPLIERS APPENDIX.*

[b]DDC, direct digital control

larly if high-cell-density processes are needed; Alternate Protocol 1), and ease of maintenance and servicing. Utilities such as air, water, steam, and electricity are needed in defined qualities and quantities. Consult the fermentor manufacturer for specific requirements.

Materials

Transformed *E. coli* host strain expressing the protein of interest (*UNIT 5.2*)
LB medium with antibiotic (*UNIT 5.2*)
Antifoam: polypropylene glycol (PPG) 2000 (pure; store indefinitely at room temperature)
Nutrient solution (optional, for fed-batch methods; see Alternate Protocol 1)
ECPM1 medium (see recipe)
85% (w/v) H_3PO_4
28% (w/v) NH_4OH
10 N NaOH

Fermentor (Fig. 5.3.1)
5-ml and 50-ml sterile syringes and 1 × 40–mm needles
Centrifuge, with capacity up to 6 liters per run (for 1- to 10-liter scale; e.g., Sorvall RC3B, equipped with H6000A rotor) or continuous feed centrifuge (for scale >10 liters; Fig. 5.3.2; e.g., Sharples AS16)
Heat-sealable plastic bags and heat-sealing apparatus

Additional reagents and equipment for inducing protein expression and for preparing samples for polyacrylamide gel analysis (*UNIT 5.2*)

Prepare the inoculum and fermentor

1. Prepare an overnight culture of the transformed *E. coli* host strain expressing the protein of interest by inoculating a suitable amount of LB medium, supplemented with the appropriate antibiotic to maintain plasmid-carrying cells, with a single colony picked from an LB/antibiotic plate or retrieved from a vial of frozen cells (see *UNIT 5.2*). The next morning, read the OD_{600} against water as a blank to verify that the

BU IASI CENTRAL UNIVERSITY LIBRARY

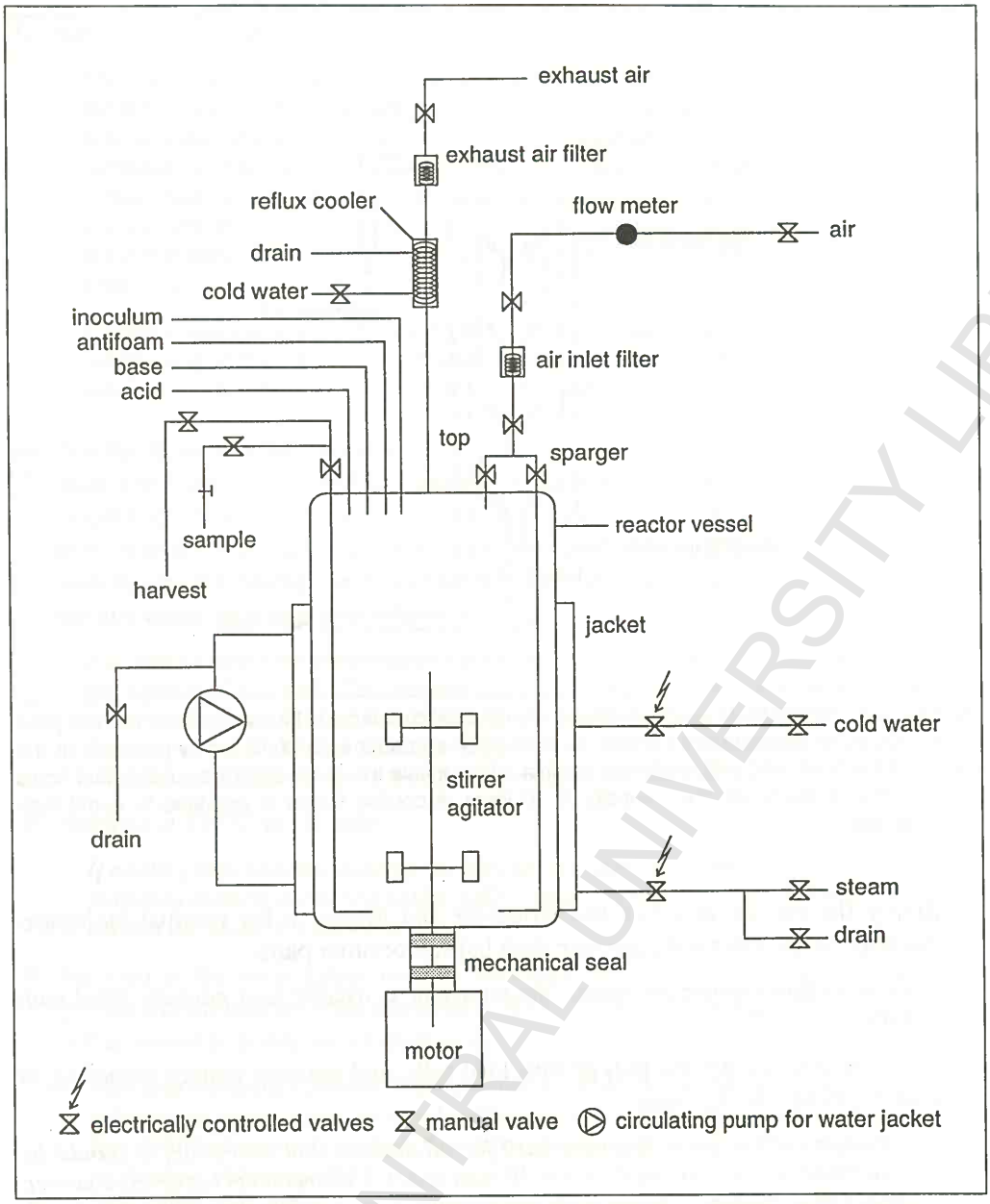


Figure 5.3.1 Schematic diagram of a bacterial fermentor, depicting major functions. Variations in the design can be made to suit specific requirements.

culture has reached stationary phase (i.e., OD>1.5, see Support Protocol 1). Use the overnight culture to inoculate the fermentor.

The size of the inoculum can be varied between 1% and 10% of the intended fermentation volume. For a 10-liter fermentor, the required inoculum of 100 to 1000 ml can be conveniently prepared in 500-ml to 5-liter shaker flasks (a maximum ratio of liquid/flask volume of 0.3 ensures that the culture is well aerated). Fermentors of sizes above 10 liters can be inoculated with several shaker flasks, representing 1% of the fermentor volume. The contents of all the inoculum flasks are pooled into a single large sterile container, fitted with the appropriate transfer line.

The choice of antibiotic is dictated by the transformation vector, and the incubation temperature depends on the expression system (use 30°C for p_L-based systems and 37°C for others). The inoculum is needed in step 15 after the fermentor has been loaded with fermentation medium and sterilized.

BOULDER CENTRAL UNIVERSITY LIBRARY

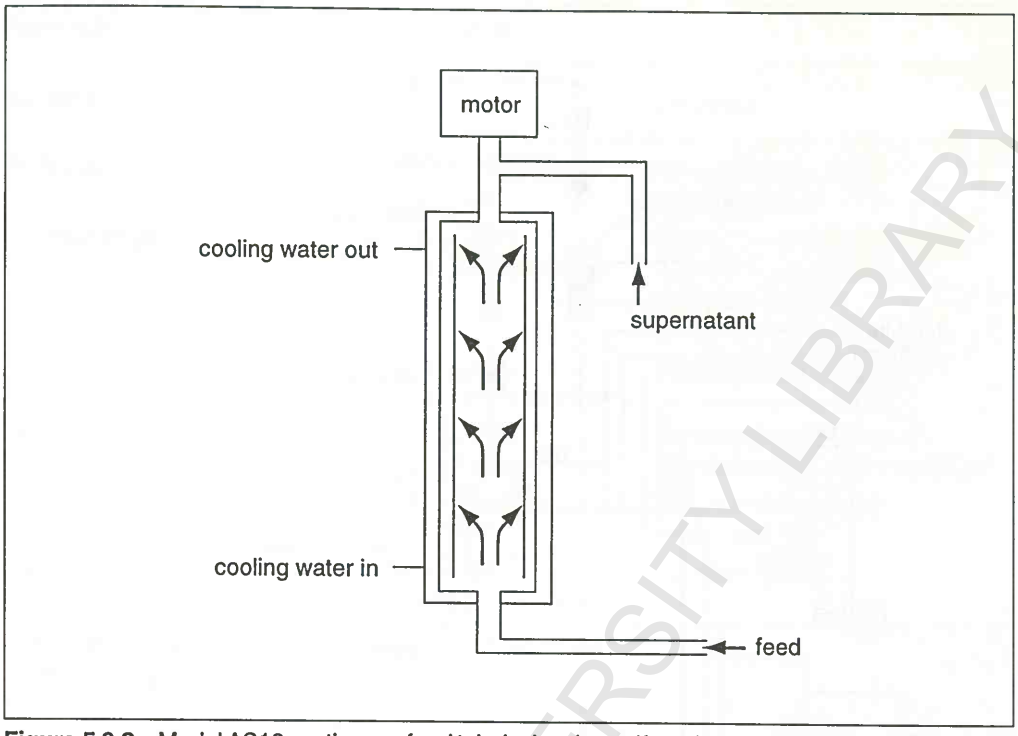


Figure 5.3.2 Model AS16 continuous feed tubular bowl centrifuge (Alfa Laval). Other models (disc stack, with pellet ejection) are suitable for pilot-scale operations. Broth is fed by pressure at the bottom of the bowl, and cells sediment against a liner inside the bowl. Cell-free supernatant flows from the top of the bowl. The capacity is ~6 liters. A cooling jacket is provided to avoid high temperatures.

2. Empty the reactor vessel of the fermentor and inspect it for residual biological material on the inner wall, agitator shaft baffles, or other parts.

   Between fermentation processes, the fermentor is usually kept partially filled with water.

3. Calibrate sensors: pH and pO_2 probes, load cells, and pressure gauges according to manufacturer's instructions.

   Although calibration is not necessary for all sensors and every run, it should be performed on a regular basis (every 10 runs or so). The pH and pO_2 probes, however, should be recalibrated before each run because of the possible drift of the signal with time or fouling of the probe membrane.

   The calibration of the pO_2 probe done here is a rough determination of whether the probe is responding to changes of dissolved O_2. The precise calibration for pO_2 is performed at step 10, after the probe is in the sterilized fermentation medium, in an environment closest to the one of the process.

4. Prepare accessory lines, transfer vessels, and ports, including septums and addition lines for inoculum, acid, base, antifoam, nutrients, and inducers. Sterilize by autoclaving. Also sterilize antifoam and, if applicable, nutrient solution, if possible while linked to the respective lines.

   Some nutrient solutions (e.g., for fed-batch processes) contain heat-sensitive components and should be sterilized according to the recipe (see Reagents and Solutions).

5. Prepare ECPM1 medium by mixing all heat-sterilizable ingredients with water to ~80% of the final volume. Pour into fermentor.

   For fermentation volumes exceeding 5 liters, dissolve all ingredients as a concentrate (at 100× the final required concentration) and pour into the fermentor.

BJU IASI CENTRAL UNIVERSITY LIBRARY

6. Adjust presterilization volume by adding water directly to the reactor vessel.

The presterilization volume should be calculated from the final culture volume (which should not exceed ¾ of the total fermentor volume) by deducting volumes of inoculum, heat-sensitive ingredients and antibiotic to be added and water that evaporates during sterilization. The addition of water resulting from steam condensation during sterilization must also be taken into account. The amounts of liquid loss or gain should not exceed 10% of the final culture volume and is best empirically determined by measuring accurately the liquid level pre- and post-sterilization in a mock trial sterilization with water.

Culture volume does vary during the process, due to mass cell accumulation, sampling, additions of inducers and for pH and foam correction, but these phenomena usually result in a net volume variation that is negligible (except for fed-batch processes).

Sterilize the fermentor and medium

7. Start sterilization cycle. For autoclavable glass transfer vessels (usually 1 to 5 liters), disconnect all cables and tubings from the fermentor; allow the inside of the chamber to be ventilated through one or two lines (protected with air filters) that communicate with the reactor headspace. For transfer vessels that can be sterilized in situ, engage the automatic sequence controller.

It is best to insert the temperature sensor of the autoclave into the temperature port of the reactor, if possible. This ensures that the autoclave control system will sense the exact temperature of the broth inside the reactor. For in situ sterilizable vessels, most fermentors are fitted with an automatic sequence controller that initiates and controls the whole sterilization cycle with predetermined parameters.

8. Sterilize at 121°C for 30 min.

If cooling proceeds too quickly, an internal vacuum may form that will open the way to contamination. A cooling rate of ≤40°C/min is acceptable—if it is higher than this, restrict the cooling water flux.

9. As soon as the temperature reaches the range of 102° to 108°C, apply headspace pressure to the reactor with sterile air (0.05 to 0.1 bar). Start water flow through the reflux cooler to condense exhaust gases.

Applying headspace pressure is important to avoid an internal vacuum and to create a barrier to any potential external contaminant. It is also good practice, once pressure has stabilized, to set a small air flow through the headspace (0.2 liters air/culture/min) to dry the exhaust gas filter.

Conduct final preparations for inoculation

10. For pH correction, pour 85% H_3PO_4 into one transfer vessel (acid) and 28% NH_4OH into the other (base) under a biological hood. Connect acid, base, and antifoam lines aseptically to the fermentor. Once the temperature is ≤40°C, proceed to step 11.

Aseptic connections are easily made by using a sterile needle to puncture a sterile septum port. Use flame sterilization and/or a 70% ethanol spray to ensure sterility. Some connectors can be sterilized again by steam after the connection is completed.

11. Add heat-sensitive ECPM1 medium ingredients (the 1 M $MgCl_2 \cdot 6H_2O$ previously sterilized by 0.22-µm filtration; see Reagents and Solutions) with a sterile syringe by puncturing a septum port with the needle.

12. Calibrate pO_2 sensor by sparging air at 0.5 vol gas/vol culture/min with vigorous agitation (≥500 rpm) and adjusting the 100% saturation point.

The zero point of the sensor can be calibrated by sparging N_2 gas instead of air, or with a sensor simulator, but this point should not need frequent recalibration—after every 10 sterilization cycles should suffice.

BOKU - IAM - UNIVERSITÄTSBIBLIOTHEK

13. Adjust all environmental parameters to the desired set points.

Typical parameter settings are: temperature, 30° or 37°C; pH, 7.0; agitation, 300 rpm; air flow, 1 vol/vol culture/min; pO_2, 100% air saturation.

14. Withdraw a small sample aseptically and check for sterility (see Support Protocol 2).

A sterile fermentor can be prepared 2 to 3 days in advance. If this is done, it is best to recheck the sterility of a fresh sample taken just before inoculation. Because the results of this sterility check are not known for another 12 to 16 hr, the procedure is initiated and continued. In case of a contamination revealed after inoculation, the batch should be considered suspect and/or discarded. A massive contamination (which could occur if the fermentor has been prepared more than 24 hr in advance) will be obvious from the evolution of the pH and/or pO_2 and turbidity of the sample. If any contamination is detected before inoculation, the entire run should be cancelled and the fermentor cleaned (see step 24).

Inoculate the fermentor, growth cells, and induce protein expression

15. Aseptically connect inoculum line to fermentor septum port. Inoculate fermentor by transferring the inoculum (step 1) aseptically into the reactor vessel. Disconnect inoculum line.

A short spray of 70% ethanol on the septum and needle improves sterility.

Inoculation is achieved either via an in-line peristaltic pump or by pressurizing the inoculum vessel with sterile air.

16. Monitor growth by measuring OD_{600} of samples taken at 1-hr intervals (see Support Protocol 1). When the target cell concentration is reached, proceed to induction.

Target cell concentration depends on the strain and expression system. As a general rule, and in the absence of any previous experience with a given strain, a target of OD 5 to 10 is applicable. It is possible, however, that the same process can also be run with induction at a later stage (OD 15 to 20), but this needs empirical determination.

17. Induce protein expression via temperature or chemical induction using one of the following methods, as appropriate.

    a. *For the p_L system:* Perform a fast temperature shift (within a few minutes) from 30° to 42°C and incubate 5 hr.

    b. *For the trp* promoter system: Add IAA to a final concentration of 25 mg/liter (see *UNIT 5.2*). Incubate to stationary phase (usually overnight).

    c. *For the lac and tac promoter and T7 systems:* Add IPTG to a final concentration of 0.5 to 1.0 mM (*UNIT 5.2*) and incubate 5 hr.

18. At 1-hr intervals, withdraw samples of the culture, monitor growth by OD_{600} measurements, and prepare them for polyacrylamide gel analysis, as described in *UNIT 5.2*, to monitor recombinant protein expression.

A band of the expected molecular weight should appear in lanes of induced samples, but not in those of uninduced or early-time-of-induction samples. Results of this analysis cannot be known before the projected harvesting time, so for the first experiment, the process should be run with the time lines indicated above. The best combination of high expression (judged by band intensity on the gel) and high cell density should dictate optimal harvesting time—these conditions can then be applied to repeat experiments.

It it also important to monitor growth closely during the induction phase because some strains rapidly lyse. In case of cell lysis, proceed to harvesting.

PSU IASL / CENTRAL UNIVERSITY LIBRARY

Harvest the cells

For small fermentors (1 to 10 liters):

19a. To harvest cells from small fermentors, pressurize the fermentor and transfer the reactor contents into an intermediate container.

20a. Distribute liquid culture, preferably under a biological hood, into centrifuge bottles. Record final culture volume in order to estimate final cell yield.

Aerosols should be minimized at all times. Use closed containers fitted with sterile filter vents.

21a. Centrifuge 15 min at 5000 × *g* (4000 rpm in an H6000A rotor), 4°C.

22a. Collect pellets from each centrifuge bottle and place together in a heat-sealable plastic bag. Proceed to step 23.

If the protein is secreted to the extracellular medium, collect the supernatant and concentrate by tangential flow ultrafiltration.

For large fermentors (>10 liters):

19b. For larger fermentors, prepare for cell harvest by removing all flexible tubing connections on the reactor, stopping pH and pO_2 control loops, and setting the temperature set point at minimum. Set pressure set point to 0.8 bars and direct gas flow to the headspace of the reactor.

In Figure 5.3.2, the broth is transferred from the fermentor to the bottom of a continuous feed centrifuge by pressurizing the reactor through the headspace. Consult the manufacturer's instructions for specific operating conditions. It is important to cool the culture liquid quickly to minimize proteolytic degradation.

20b. Start the continuous feed centrifuge and accelerate to an effective centrifugation force of ≥10,000 × *g*. When bowl speed is stable, open connecting valves from the fermentor to the centrifuge.

21b. Set inlet flow rate to the centrifuge at ~1 to 2 liters/min. Centrifuge at 4°C until all the medium has been processed.

22b. Stop centrifuge, dismount the bowl, and collect the pelleted cells into a heat-sealable plastic bag.

If the purification process cannot handle the entire pellet volume, divide the material into smaller aliquots.

If the protein is secreted to the extracellular medium, collect the supernatant and concentrate by tangential flow ultrafiltration.

Store the harvested cells and clean the fermentor

23. Retain a small amount of the pellet (50 to 100 mg) in a separate microcentrifuge tube for polyacrylamide gel analysis (step 18) to verify that centrifugation has not altered the target protein expression level. Close pellet-containing bags by heat sealing. Record total pellet weight and freeze at –20°C.

Most recombinant proteins in the pelleted cells remain stable for months when stored at –20°C. However, some membrane receptors (e.g., neurokinin, adrenergic, transmembrane receptors in general) are best preserved by snap freezing in dry ice/isopropanol and storing at –70°C.

24. To clean the fermentor, fill it with water and adjust to pH 11.0 to 12.0 by adding 10 N NaOH. Set temperature to 50°C and leave with vigorous agitation overnight.

Local biosafety rules may require sterilization of the reactor before cleaning. For

BCU ARCHIVE CENTRAL UNIVERSITY LIBRARY

sterilization, fill the reactor with water and proceed to steps 7 and 8. If the fermentor is still not clean after this treatment, alternate alkali and acid (H_3PO_4) treatments.

25. Empty the reactor, rinse extensively with water, and sterilize as described in steps 7 and 8. Prepare the fermentor for another run, if desired.

Electrochemical sensors (pH, pO_2, pCO_2) should be removed and stored according to manufacturer's instructions. The fermentor should then be filled with water, sterilized, and stored sterile until the next run.

ALTERNATE PROTOCOL 1

PROTEIN PRODUCTION IN HIGH-CELL-DENSITY PROCESSES: FED-BATCH METHODS

Fed-batch fermentations take advantage of the high-protein-synthesizing capacity of an exponentially growing culture while minimizing the buildup of toxic by-products by keeping growth carbon source–limited. This protocol describes a simple glycerol-supplementation method that is applicable to a wide range of culture conditions. Cells are first grown in a batch fermentation in a medium containing 10 g/liter glycerol. After the initial glycerol is consumed, the culture is supplemented with an increasing feed rate of concentrated medium containing 900 g/liter glycerol. Except for cell density, induction of protein synthesis and harvest are the same as for batch fermentation (see Basic Protocol 1).

Additional Materials (also see Basic Protocol 1)

HCDF1 medium (see recipe)
HCDM1 medium (see recipe), containing 10 g/liter glycerol
O_2 gas, pure

1. Prepare an inoculum culture of the transformed *E. coli* host strain expressing the protein of interest, then set up the fermentor as for batch fermentation (see Basic Protocol 1, steps 1 to 4).

2. Prepare HCDF1 feed medium in a feed tank (transfer vessel) and set on a balance or a load cell to keep track of feed consumption.

3. Proceed to preparing the fermentation medium, sterilizing the medium and fermentor, and conducting the final preparations for inoculation for batch fermentation procedure (see Basic Protocol 1, steps 5 to 14), substituting HCDM1 batch medium containing 10 g/liter glycerol as the fermentation medium.

4. Inoculate the fermentor and grow cells (see Basic Protocol 1, steps 15 and 16). After the initial glycerol is consumed, start the feed of HCDF1 medium.

A sharp rise in pO_2 is indicative of reduced oxygen consumption and thus of glycerol becoming limiting. Monitor growth during this phase to estimate maximum growth rate (see Sinclair and Cantero, 1990, for methodology).

5. Increase rate of addition of HCDF1 medium gradually and withdraw samples to follow the increase in OD_{600}. Adjust rate of HCDF1 medium addition to the target feed rate. Monitor feed rates by measuring the mass of the feed tank over time.

The target feed rate can be estimated as

$$F = \mu \times Vf \times OD/(Y \times 0.4 \times 0.9)$$

where F is the feed rate (ml/hr), μ is the target growth rate (hr^{-1}), Vf is the fermentor volume (liters), OD is the optical density at 600 nm, Y is the yield of cell mass per unit carbon source (g/g), 0.4 represents the conversion factor from OD to cell mass (OD/g/liter), and 0.9 is the glycerol content in the feed solution (g/ml). An estimate of

BU IASI / CENTRAL UNIVERSITY LIBRARY

80% of the maximum growth rate (step 4) is a good starting value. An estimate of 0.3 is a good value for the yield, Y.

When monitoring feed rates, make sure to take into account the density of the feed and the increase in fermentation volume with feeding. The target growth rate is dependent on each strain and each process. To avoid accumulation of toxic by-products (organic acids), the target growth rate should be set below 80% of the maximum growth rate observed during the batch phase.

6. Blend pure O_2 gas into the air that is sparged into the reactor (either by solenoid valve pulses or by a motor-driven proportional valve) to make sure that pO_2 is properly controlled (set point between 20% and 50%).
7. Once the target high biomass density is obtained (OD_{600} above 50), induce expression of the protein as in batch fermentation procedure (see Basic Protocol 1, step 17).

   Final cell densities can be as high as 100 to 150 OD units.

   Induction by a chemical inducer can also be achieved by simply adding the inducer to the feed solution (gradual induction).
8. Proceed to monitoring expression and harvesting (see Basic Protocol 1, steps 18 to 23) while continuing to match the rate of HCDF1 supplementation with growth.

IN VIVO LABELING OF RECOMBINANT PROTEINS WITH HEAVY METAL DERIVATIVES

ALTERNATE PROTOCOL 2

To maximize incorporation of selenomethionine in place of methionine in recombinant proteins to facilitate structural studies, it is preferable to use a methionine auxotroph host strain. The protocol is best suited for small fermentors (1 to 10 liters) because the medium change step requires sterile harvesting via a batch pot centrifuge.

Additional Materials (also see Basic Protocol 1)

Methionine auxotroph host strain: *E. coli* DL41 or B834 (Table 5.2.1)
Expression vector containing gene for the protein of interest
DLM medium (see recipe)
Complete rich medium (e.g., LB medium; *UNIT 5.2*)

Additional reagents and materials for electroporation and testing expression of transformants (*UNIT 5.2*)

Construct the expression strain

1. Transform methionine auxotroph host strain with expression vector containing the gene for the protein of interest (see *UNIT 5.2*).

   For T7 expression systems, use E. coli B834 (with or without pLys plasmid) as a host (also see UNIT 5.2 Commentary and see Table 5.2.1).
2. Test expression of transformants in a shaker flask induction test (see *UNIT 5.2*).

Set up the fermentation and induce protein expression

3. Prepare DLM medium (same volume as fermentor) and keep at room temperature.
4. Prepare fermentor with complete rich medium and inoculate with the recombinant strain (see Basic Protocol 1, steps 1 to 16).
5. Grow overnight in rich medium at 30°C for p_L strains, 37°C for others.

B. C. U. „M. [illegible] IAŞI
C H [illegible]

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

6. Withdraw sample and measure OD_{600}. Harvest cells under sterile conditions by batch pot centrifugation (see Basic Protocol 1, steps 19a to 22a). Stop all control loops on the fermentor.

   OD values should be in the range of 20 to 30 after overnight growth. Harvest only part of the culture. Calculate the volume to be harvested so that, once the cells are resuspended in the same volume of DLM labeling medium, the OD_{600} will be between 10 and 15.

7. Resuspend the pellet in DLM medium. Transfer to a sterile container and then back to the fermentor. Resume all control loops.

   In practice, there will be several pellets, distributed in several centrifuge bottles. Pellets are individually resuspended, then suspensions are pooled into a single vessel to avoid having to transfer each separate bottle back into the fermentor. Remember that sterility is a must here, and handling one container is less risky than dealing with several.

8. Wait for 30 min until all environmental parameters are back to their set points and are stable before inducing expression (see Basic Protocol 1, step 17). Proceed to monitoring expression and harvesting cells (see Basic Protocol 1, steps 18 to 23).

ALTERNATE PROTOCOL 3

STABLE ISOTOPIC LABELING OF RECOMBINANT PROTEINS

Uniform incorporation of stable isotopes in recombinant proteins is an alternative to selective labeling of a given amino acid (Alternate Protocol 2) to facilitate structural studies. The method requires the use of a semi-defined medium containing salts, trace elements, and yeast extract. Cells are supplied with [^{13}C]glucose, $^{15}NH_4Cl$, and/or D_2O, and induced to synthesize recombinant protein labeled with ^{13}C, ^{15}N, and/or ^{2}H. The three labels can be used individually or in combination. This protocol makes use of the fed-batch method (Alternate Protocol 1) but substitutes (labeled) glucose for glycerol as the carbon source and is run at much lower cell densities.

Additional Materials (also see Basic Protocol 1)

Transformed *E. coli* host strain expressing the protein of interest (*UNIT 5.2*)
HCDM1 medium (see recipe)
50% (w/v) [^{13}C]glucose
100 g/liter $^{15}NH_4Cl$ stock solution
D_2O
Base for pH correction: 2 N NaOH (for ^{15}N label)

1. Grow an inoculum culture of the transformed *E. coli* host strain expressing the protein of interest, then prepare fermentor and start the batch fermentation with HCDM1 medium (see Basic Protocol 1, steps 2 to 14), setting up 50% [^{13}C] glucose as the nutrient solution (place feed tank on a balance or load cell).

   In this protocol, the carbon source is labeled glucose instead of glycerol.

   For ^{2}H labeling, use D_2O instead of H_2O in making up HCDM1 medium.

2. After medium sterilization add 10 g/l [^{13}C]glucose and the first 5-ml portion of 100 g/liter $^{15}NH_4Cl$ stock solution per liter fermentation volume (using a sterile syringe; see Basic Protocol 1, step 12).

   Glucose is added after sterilization to avoid partial destruction when heat sterilized with nitrogen-containing compounds present in the medium. To achieve the target 10 g/liter concentration, add 40 ml of the feed solution per liter of culture.

BU IASI / CEN / PAL - UNIVERSITY LIBRARY

$^{15}NH_4Cl$ is added in three separate batches to avoid inhibition by growth by excess NH_4Cl in the medium.

3. Inoculate fermentor (see Basic Protocol 1, step 15). Allow cells to grow. When the initial [^{13}C]glucose is consumed, start feed with [^{13}C]glucose.

   A sharp rise in pO_2 is indicative of glucose becoming limiting and thereby reducing oxygen consumption. To set initial and subsequent feed rates, use the equation in Alternate Protocol 1, step 5, substituting 0.5 for 0.9 as the glucose feed content (g/ml), using a target growth rate of 0.3/hr and an estimate of 0.4 for the yield Y.

4. Monitor growth by withdrawing samples periodically and measuring OD_{600}. At an OD_{600} ~5, add second batch of $^{15}NH_4Cl$ and induce protein expression (see Basic Protocol 1, step 17).

   In this protocol a high cell density is not desirable, and therefore no pure O_2 should be necessary for pO_2 control.

5. At 1 hr postinduction add the last batch of $^{15}NH_4Cl$.

6. Proceed to monitoring protein expression and harvesting (see Basic Protocol 1, steps 18 to 23) while continuing to match the rate of glucose supplementation with growth.

GROWTH MONITORING

SUPPORT PROTOCOL 1

Bacterial growth can be monitored off-line by several different methods (Lech and Brent, 1994a). For fermentation process control, spectrophotometry is fast, reliable, and simple. Microscope counting is an alternative but is more cumbersome and time-consuming. On-line sensors also exist and have been used successfully in some cases (Schügerl et al., 1993). Most sensors, however, are sensitive to particulate matter, gas bubbles, fouling, and calibration drifts.

Materials

Test sample (e.g., *E. coli* inoculum or fermentation medium)
Culture medium (optional, for diluting visibly turbid cultures)

Spectrophotometry cuvettes (e.g., 1-cm light path)
Fixed-wavelength spectrophotometer (e.g., 600 or 660 nm)
Count slide *or* hemacytometer
Phase-contrast microscope (400×)

For spectrophotometry:

1a. Place test sample in spectrophotometry cuvette. Fill reference cuvette with water.

2a. Transfer cuvettes to a fixed-wavelength spectrophotometer and measure the optical density at 600 nm.

   The level of absorbance at 600 nm will depend on the distance between the cuvette and the detector and will vary among instruments, often by a factor of 2. It is thus wise to calibrate each instrument by recording the OD_{600} of a culture that contains a known number of cells determined by some other method, such as observation on a count slide or titering for viable colonies (Lech and Brent, 1994b).

   For a culture grown in rich medium, a good rule of thumb is that each 1 OD unit is roughly equivalent to 10^9 cells/ml. OD readings above 0.4 cannot be accurately correlated with cell numbers and the sample should be diluted 10- or 100-fold with PBS to bring reading back into the 0.4 range.

BOWMAN / CENTRAL UNIVERSITY LIBRARY

For cell counting:

1b. Take a clean count slide (or hemacytometer) and cover it with a clean coverslip. Dip a 0.1- or 1-ml pipet into the test sample, allow a small drop of liquid to form on the end of the pipet, and touch it lightly to the surface of the slide at the periphery of the coverslip.

2b. Put the slide on the stage of a phase-contrast microscope set to 100×, and focus on the cells. Calculate cell densities.

Each cell in a small square of a count slide is equivalent to 2×10^7 cells/ml.

SUPPORT PROTOCOL 2

STERILITY CHECKING

Although results are known only after an overnight incubation, checking the sterility of the fermentation medium prior to inoculation is essential to ensure reproducible results and aids in troubleshooting potential problems. To absolutely eliminate contamination during the test itself, it is preferable to perform this protocol under a microbiological hood. A quick check of sterility requires a sample size of 100 µl, whereas more sensitive detection requires up to 100 ml.

Materials

Test sample (e.g., fermentation medium)
LB plates with and without antibiotic (*UNIT 5.2*)

0.45-µm membrane (Millipore, type HA or HC)
Sterile filter holder

For a quick check: Spread 100 µl test sample onto one LB plate and one LB/antibiotic plate. Incubate at 30°C overnight. Inspect for appearance of any colonies.

For a more sensitive detection: Filter up to 100 ml test sample by vacuum aspiration through a 0.45-µm membrane set in a sterile filter holder. Disassemble filter holder and transfer membrane to an LB agar plate, face up. Incubate plate 24 to 48 hr at 30°C. Inspect membrane for appearance of any colonies.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

DLM medium

2.5 g adenine·HCl
2.2 g L-glycine
1.3 g guanosine
1 g L-alanine
1 g L-arginine
0.8 g L-aspartic acid
0.06 g L-cystine
1.3 g L-glutamic acid
0.66 g L-glutamine
0.12 g L-histidine
0.5 g L-isoleucine
0.5 g L-leucine
0.7 g L-lysine
0.2 g L-methionine

continued

PSU IASL CENTRAL UNIVERSITY LIBRARY

0.26 g L-phenylalanine
0.2 g L-proline
0.2 g seleno-L-methionine
4.2 g L-serine
0.46 g L-threonine
0.34 g L-tyrosine
0.46 g L-valine
0.34 g thymine
1 g uracil
20 g glucose
0.75 g NH_4Cl
1.2 g K_2HPO_4
0.4 g $MgCl_2 \cdot 6H_2O$
10 ml trace element solution 1 (see recipe)
H_2O to 1 liter

CAUTION: Seleno-L-methionine, which replaces methionine, is highly toxic.

Dissolve each ingredient one after the other in water, except for L-aspartate, L-glutamate, L-tyrosine, guanosine, thymine, and uracil, which need to be dissolved first in a small amount of 0.1 N NaOH.

Sterilize by filtration through a membrane with 0.22-µm pores. Store at room temperature.

ECPM1 medium

4 g K_2HPO_4
1 g KH_2PO_4
1 g NH_4Cl
2.4 g K_2SO_4
132 mg $CaCl_2 \cdot 2H_2O$
10 ml trace element solution 1 (see recipe)
20 g Casamino Acids (enzymatic hydrolysate; Difco)
3 g yeast extract (Difco)
40 g glycerol
H_2O to 800 ml

Mix ingredients, pour into fermentor, and sterilize by heat (see Basic Protocol 1, steps 5 to 8) adding balance of water to 1 liter in step 6. Add 2 ml of 1 M $MgCl_2 \cdot 6H_2O$ (sterilized by filtration through a membrane with 0.22-µm pores) as directed in step 12 of Basic Protocol 1. Once sterilized, the medium can be stored up to 1 week at 4°C.

HCDF1 medium

900 g glycerol
H_2O to 910 ml
Sterilize by heat
Add 20 ml 1 M $MgCl_2 \cdot 6H_2O$ (sterilized by filtration through 0.22-µm membrane)
Add 70 ml trace element solution 2 (see recipe)

BCU IASRI / CENTRAL UNIVERSITY LIBRARY

HCDM1 medium

10 g K_2HPO_4
10 g KH_2PO_4
9 g $Na_2HPO_4 \cdot 2H_2O$
1.1 g NH_4Cl
9 g K_2SO_4
10 ml trace element solution 2 (see recipe)
2 g yeast extract
H_2O to 800 ml

Sterilize by heat (see Basic Protocol 1, steps 5 to 8), adding balance of water to 1 liter in step 6. Add 2 ml 1 M $MgCl_2 \cdot 6H_2O$ (sterilized by filtration through 0.22-μm membrane; see Basic Protocol 1, step 12). Supplement with carbon source (glycerol or glucose) as indicated in Alternate Protocol 1 and Alternate Protocol 3.

Trace element solution 1

5 g EDTA
0.5 g $FeCl_3 \cdot 6H_2O$
0.05 g ZnO
0.01 g $CuCl_2 \cdot 2H_2O$
0.01 g $Co(NO_3)_2 \cdot 6H_2O$
0.01 g $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$

Dissolve EDTA in 700 ml water. Dissolve each element separately by adding just enough 10 N HCl to bring it into solution. Add each element in solution one after the other to the EDTA solution. Add water to 1 liter. Adjust pH to 7.0 with concentrated NaOH. Autoclave. Store in the dark at 4°C.

Trace element solution 2

5 g EDTA
6 g $CaCl_2 \cdot 2H_2O$
6 g $FeSO_4 \cdot 7H_2O$
1.15 g $MnCl_2 \cdot 4H_2O$
0.8 g $CoCl_2 \cdot 6H_2O$
0.7 g $ZnSO_4 \cdot 7H_2O$
0.3 g $CuCl_2 \cdot 2H_2O$
0.02 g H_3BO_3
0.25 g $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$

Dissolve EDTA in 700 ml water. Dissolve remaining ingredients separately in just enough 10 N HCl to achieve solution and add one after the other to the EDTA solution. Add water to 1 liter. Adjust to pH 7.0 with concentrated NaOH. Autoclave. Store in the dark at 4°C.

COMMENTARY

Background Information

The choice of a strategy for fermentation is based primarily on the type of expression system (see UNIT 5.1). For example, chemical induction (e.g., via *trp*, *lac*, or tac promoters) may be more feasible than temperature induction (e.g., via the lambda promoter) in larger-volume fermentors. The strategy also depends on the final intended use for the protein: immunization and generation of antibodies, structure-function investigations, structural studies, or in vivo assays.

Because the fermentation process (i.e., the media, temperature, pH, and pO_2 used) can affect the solubility and localization of the protein, it should be designed with the purification process in mind (Chapter 6).

Localization of the protein in the *E. coli* cell (cytoplasmic, periplasmic, or secreted in the extracellular medium) is also a prime factor in

BAU IAST / CENTRAL UNIVERSITY LIBRARY

the design of the process. Finally and most importantly, the toxicity of the target protein for *E. coli* is a key determinant in the success of the production. For example, cell lysis has been observed repeatedly when over-expressing proteins of the apolipoprotein family.

Fermentation of recombinant *E. coli* has been developed by many different research groups, and several well-documented processes have been published (Knorre et al., 1991; Strandberg et al., 1991; Yang, 1992; Gram et al., 1994). In general, accurate control of environmental parameters (e.g., temperature, pH, and pO_2) is prerequisite for any production. This will also ensure reproducibility among different batches. Table 5.3.1 summarizes vendors of fermentation equipment that we have used successfully. The list is not meant to be exhaustive, and local supply may vary.

Although most recombinant protein production protocols described in the literature are based on a simple batch type of operation (see Basic Protocol 1), fed-batch methods (see Alternate Protocol 1) allow a more productive use of a given reactor by allowing growth of the culture to a high cell density prior to induction. The principle is to supply increasing amounts of nutrients so as to match exponential growth of the culture while minimizing accumulation of toxic by-products. Growth is made carbon source–limited in order to keep the concentration of residual carbon source low and to maximize yield of recombinant proteins.

Fermentations may also be designed to produce labeled products. The replacement of methionine residues by selenomethionine (SeMet) in proteins (see Alternate Protocol 2) was first suggested as a new strategy to tackle a major hurdle in protein X-ray crystallography, namely, determination of the phases (Yang et al., 1990). The approach consists of using crystals of the SeMet-substituted protein in multiwavelength anomalous diffraction (MAD) experiments, which require access to a synchrotron. The SeMet protein can also be used, however, in fixed-wavelength analysis to provide landmarks (i.e., the heavy atoms of selenium) that facilitate data analysis and resolution of the structure.

Nuclear magnetic resonance (NMR) spectroscopy is used increasingly for determining the three-dimensional structure of proteins (up to 40 kDa) in solution. In conventional NMR techniques, however, proteins, because of their large size, generate spectra with overlapping signals, in particular in the region of $_1H$ resonances. This led several teams to develop a general method for labeling proteins with different stable isotopes such as ^2H, ^{13}C, and/or ^{15}N (see Alternate Protocol 3). Isotopic labeling allows NMR specialists to study heteronuclear coupling and to use powerful two-, three-, and four-dimensional editing of spectra. These techniques greatly facilitate assignment of resonances (for a review, see Clore and Gronenborn, 1994).

Critical Parameters and Troubleshooting

Equipment. For optimal operation, several fermentor characteristics are essential. Heat transfer is best achieved through a double jacket around a stainless steel base. Water flow should be guided in loops around the vessel wall to avoid any dead zones. Steam, if available, is the most efficient means of heating. Heating efficiency is critical for temperature shift inductions, where the broth should be heated from 30° to 42°C in a few minutes. In fact, for fermentors with a working volume >10 liters, there should be provision for direct steam injection to obtain a very quick temperature shift. Cooling requires consumption of cold water, which, in some instances can be recycled. Recycling is most useful for installations where supercold water (–4° to 4°C) is used.

In addition to the essential sensors (i.e., temperature, agitation, dissolved O_2, pH, and foam level), sensors such as gas flow, pressure, load, redox potential, culture fluorescence, and mass spectrometer on exhaust gas (Todd, 1989) can be mounted to monitor the environment. Halling (1990) gives detailed information on basic principles and calibration of the essential monitors and Dusseljee and Feijen (1990) review sensors and process control information.

Dissolved oxygen and O_2 transfer can be controlled by agitation, air flow, gas composition, and headspace pressure. In practice, using high-agitation-rate motors with Rushton turbines and sparging O_2-enriched air allow maximal transfers, but oxygen transfer can still become limiting at high cell densities. The average specific oxygen uptake rate for *E. coli* strains growing at a rate of 0.3 hr–1 in ECPM1 medium at 30°C is around 0.4 g O_2/g cells/hr.

Sterilization is a critical operation that can be hazardous if not properly handled. It should always be carried out according to the instructions of the fermentor supplier. An automatic controller is helpful and considerably limits the risks; it also ensures reproducibility of the operation from one batch to another. Sterility is, of course, a prerequisite for any fermentation

BOURNEMOUTH UNIVERSITY LIBRARY

experiment. Fermentor designs should be checked carefully to ensure that all parts in contact with the culture are properly sterilized by heat. The bottom region of the reactor vessel where the agitator shaft enters the fermentor is critical. Double mechanical seals with lubrication by steam condensates provide an efficient sterility barrier. Magnetically coupled agitation systems are even safer because all seals are static.

Computer control is not absolutely required for most routine purposes. It is helpful, however, for fed-batch processes and for those requiring model-based feedback control from off-line analysis. It also allows extensive data logging and analysis both from the fermentor itself and from peripheral instrumentation (e.g., off-gas analysis, flow injection analyzers).

High-cell-density methods. Fed-batch methods can be classified into two types: (i) growth-linked feeding using process parameters such as pH (Bauer and White, 1976) or pO_2 (Mori et al., 1979) to control the feed of nutrients, and (ii) model-dependent feeding, controlled by a mathematical model of growth (Yee and Blanch, 1992). The first class requires little modification to a batch operation, whereas the second class necessitates on-line computer control. In the latter case, one needs to estimate growth from off-gas measurement (Reiling et al., 1985) or by real-time measurement of growth by spectrophotometry. The feed rate can be adjusted via computer control to match the increase in growth monitored by OD_{600}. As high cell densities are obtained, limitations eventually arise from broth viscosity, oxygen transfer, or the cooling capacity of the fermentor. Figure 5.3.3 illustrates a typical fed-batch process run according to Alternate Protocol 1, with a recombinant *E. coli* strain and without induction of expression.

Isotopic labeling. As an alternative to the defined medium described in Alternate Protocol 3, a commercial medium (Celtone) is available from Martek with labeled amino acids derived from unicellular algae. Celtone gives higher cell densities and is inexpensive and easy to make. To use this medium, follow Basic Protocol 1 for growth and induction, use 5 N NaOH for pH control when labeling with ^{15}N, and add another 1× concentration of Celtone medium (100 ml of 10× stock solution per liter fermentation volume) at the start of induction.

Harvesting. Harvesting can be achieved by tangential flow microfiltration as an alternative to continuous centrifugation (Bailey et al., 1990). Tangential flow microfiltration allows the operation to be more contained, but requires a final centrifugation to obtain a cell cake. However, it can also be coupled to cell wash, diafiltration, buffer exchange, and cell breakage by continuous-flow high-pressure homogenizers.

The most common problem found in scaling up from shaker flask experiments is the decrease in expression level. In this case, one can vary the time of induction, check plasmid stability (*UNIT 5.2*), and test other host/vector combinations. Changing the composition of the medium may also solve the problem but is lengthy and labor-intensive. The absence of cell lysis should be confirmed by frequent microscopic observations, and cells must be quickly harvested as soon as lysis begins.

Sterility check. Because the results of this test are not known for 12 to 16 hr, the procedure is initiated and continued. In case of a contamination revealed after inoculation, the batch should be considered suspect and/or discarded. A massive contamination (which could occur if the fermentor has been prepared more than 24 hr in advance) will be obvious from the evolution of the pH and/or pO_2 and turbidity of the sample. If any contamination is detected before inoculation, the entire run should be cancelled and the fermentor cleaned (see Basic Protocol 1, step 24).

Anticipated Results

With an average level of expression (e.g., 10% to 15% of total cellular proteins), a normal batch process will yield ~100 g wet cell pellet per liter, containing ~500 mg of the recombinant protein. This shows that even small-scale fermentations (1 to 10 liters) are quite effective in providing the starting material for purification developments. For actual production, a fed-batch process is more typically used—this maximizes fermentor output, but also requires more manpower to operate. In this case, overall yields of 1.5 to 2 g recombinant protein per liter culture, in the crude extract, are expected.

Time Considerations

Preparation of a fermentor and its ancillaries can be done in 4 to 5 hr. Preparation of the inoculum can be started after that, allowing an overnight test for reactor sterility. The following morning inoculation can proceed, and the fermentation should be completed within 24 to 36 hr. Harvesting takes 1 to 2 hr. Analysis requires an additional 3 hr. While running the analysis, the fermentor can be cleaned, sterilized with water, and made ready for another run. Overall, a complete cycle takes 2.5 to 3 days.

BOU IAS/CERN UNIVERSITY LIBRARY

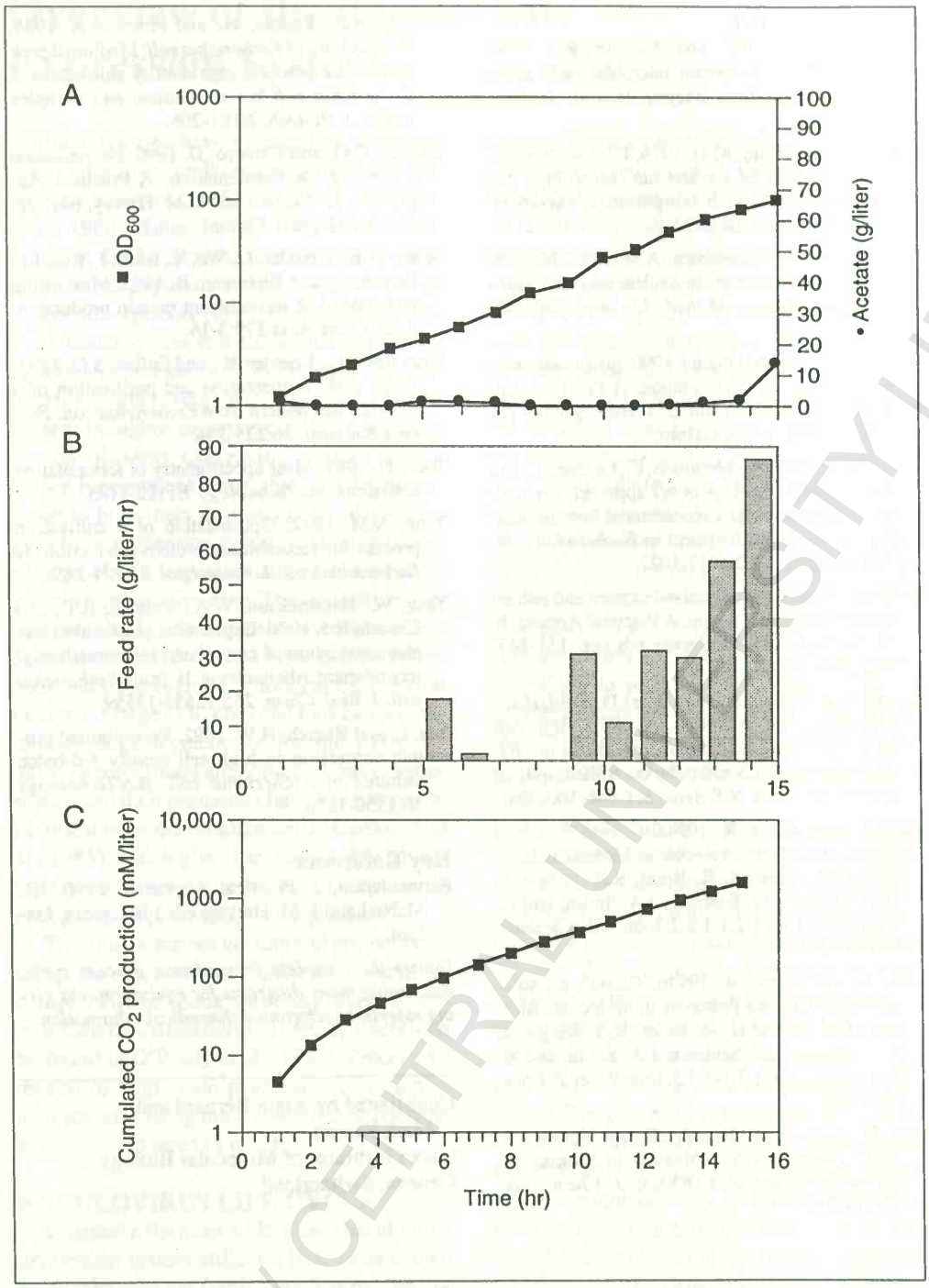


Figure 5.3.3 High-cell-density growth. (**A**) Kinetics of growth (OD_{600}) and acetate production. (**B**) Feed rate profile. Note that the feed rate was adjusted according to growth rate as described in Alternate Protocol 1. (**C**) Kinetics of cumulated CO_2 production. Mass spectrometry of the exhaust gas was performed as described by Todd (1989). Acetate formation was measured as described by Reiling et al. (1985).

BCU IAS / CENTRAL UNIVERSITY LIBRARY

Literature Cited

Bailey, F.J., Warf, R.T., and Maigetter, R.Z. 1990. Harvesting recombinant microbial cells using crossflow filtration. *Enzyme Microb. Technol.* 12:647-652.

Bauer, S. and White, M.D. 1976. Pilot scale exponential growth of *Escherichia coli* to high cell concentration with temperature variation. *Biotech. Bioeng.* 18:839-846.

Clore, G.M. and Gronenborn, A.M. 1994. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* 239:349-363.

Dusseljee, P.J.B. and Feijen J. 1990. Instrumentation and control. *In* Fermentation, A Practical Approach. (B. McNeil and L.M. Harvey, eds.) pp. 149-171. IRL press, Oxford.

Gram, H., Ramage, P., Memmert, K., Gamse, R., and Kocher, H.P. 1994. A novel approach for high level production of a recombinant human parathyroid hormone fragment in *Escherichia coli. Bio/Technology* 12:1017-1023.

Halling, P.J. 1990. pH, dissolved oxygen and related sensors. *In* Fermentation, A Practical Approach. (B. McNeil and L.M. Harvey, eds.) pp. 131-147. IRL press, Oxford.

Knorre, W.A., Deckwer, W.D., Korz D., Pohl, H.D., Riesenberg, D., and Ross, A. 1991. High cell density fermentation of recombinant *Escherichia coli* with computer controlled optimal growth rate. *Ann. N.Y. Acad. Sci.* 646:300-306.

Lech, K. and Brent, R. 1994a. Growth in liquid media. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 1.2.1-1.2.2. John Wiley & Sons, New York.

Lech, K. and Brent, R. 1994b. Growth on solid media. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 1.3.1-1.3.5. John Wiley & Sons, New York.

Mori, H., Yano, T., Kobayashi, T., and Shimizu, S. 1979. High density cultivation of biomass in fed-batch system with DO-stat. *J. Chem. Eng.* 12:313-319.

Reiling, H.E., Laurila, H., and Fiechter, A. 1985. Mass culture of *Escherichia coli*: Medium development for low and high density cultivation of *Escherichia coli* B/r in minimal and complex media. *J. Biotech.* 2:191-206.

Sinclair, C.G. and Cantero, D. 1990. Fermentation modelling. *In* Fermentation, A Practical Approach. (B. McNeil and L.M. Harvey, eds) pp. 65-85. IRL press, Oxford.

Schügerl, K., Brandes, L., Wu, X., Bode, J., Ree, J.I., Brandt, J., and Hitzmann, B. 1993. Monitoring and control of recombinant protein production. *Anal. Chim. Acta* 279:3-16.

Strandberg, L., Koehler, K., and Enfors, S.O. 1991. Large scale fermentation and purification of a recombinant protein from *Escherichia coli. Process Biochem.* 26:225-234.

Todd, D. 1989. Mass spectrometry of fermentation emissions. *BioTechnology* 7:1182-1183.

Yang, X.M. 1992. Optimization of a cultivation process for recombinant protein production by *Escherichia coli. J. Biotechnol.* 23:271-289.

Yang, W., Hendrickson, W.A., Kalman, E.T., and Crouch, R.J. 1990. Expression, purification and characterization of natural and selenomethionyl recombinant ribonuclease H from *Escherichia coli. J. Biol. Chem.* 265:13553-13559.

Yee, L. and Blanch, H.W. 1992. Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli. Bio/Technology* 10:1550-1556.

Key Reference

Fermentation, A Practical Approach. 1990. (B. McNeil and L.M. Harvey, eds.) IRL press, Oxford.

Covers the complete fermentation process cycle, illustrating many diagrams for equipment and giving essential background theoretical information.

Contributed by Alain Bernard and Mark Payton
Glaxo Institute for Molecular Biology
Geneva, Switzerland

BSU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of the Baculovirus Expression System

Baculoviruses have emerged as a popular system for overproducing recombinant proteins in eukaryotic cells (Luckow and Summers, 1988; Miller, 1988; Miller et al., 1986; Luckow, 1991). Several factors have contributed to this popularity. First, unlike bacterial expression systems (*UNITS 5.1-5.3*), the baculovirus-based system is a eukaryotic expression system and thus uses many of the protein modification, processing, and transport systems present in higher eukaryotic cells. In addition, the baculovirus expression system uses a helper-independent virus that can be propagated to high titers in insect cells adapted for growth in suspension cultures, making it possible to obtain large amounts of recombinant protein with relative ease. The majority of this overproduced protein remains soluble in insect cells, in contrast to the insoluble proteins often obtained from bacteria. Furthermore, the viral genome is large (130 kbp) and thus can accommodate large segments of foreign DNA. Finally, baculoviruses are noninfectious to vertebrates, and their promoters have been shown to be inactive in mammalian cells (Carbonell et al., 1985), which gives them a possible advantage over other systems when expressing oncogenes or potentially toxic proteins.

This unit describes vectors and protocols for using the baculovirus expression system. A comprehensive guide outlining most of the protocols described here (as well as others) can be found in O'Reilly et al. (1992). Procedures related to large-scale production of recombinant proteins using the baculovirus expression system are covered in *UNIT 5.5*.

BACULOVIRUS LIFE CYCLE

Currently, the most widely used baculovirus expression system utilizes a lytic virus known as *Autographa californica* nuclear polyhedrosis virus (AcMNPV; hereafter called baculovirus). This virus is the prototype of the family *Baculoviridae*. It is a large, enveloped, double-stranded DNA virus that infects arthropods. The baculovirus expression system takes advantage of some unique features of the viral life cycle (Fig. 5.4.1). See Doerfler and Bohm (1986) for a comprehensive review.

As with mammalian DNA viruses, the baculovirus life cycle is divided temporally into immediate early, early, late, and very late phases. Viruses enter the cell by adsorptive endocytosis and move to the nucleus where their DNA is released. DNA replication begins ~6 hr after infection. Replication is followed by viral assembly in the nucleus of the infected cell. Two types of viral progeny are produced during the life cycle of the virus: extracellular virus particles (nonoccluded viruses) during the late phase and polyhedra-derived virus particles (occluded viruses) during the very late phase of infection. Extracellular virus is released from the cell by budding, beginning at ~12 hr postinfection, and is produced at a logarithmic rate until 20 hr postinfection, after which production drops off. Polyhedra-derived virus, on the other hand, appears in the nucleus at ~18 hr postinfection and continues to accumulate as late as 72 hr postinfection, or until the cells lyse. Occluded viral particles are embedded in proteinaceous viral occlusions called polyhedra within the nucleus of infected cells. The polyhedrin protein (29 kDa) is the major protein component of the occlusion bodies.

The polyhedrin protein serves an important function for the survival and propagation of the virus in nature. Because baculoviruses are lytic, they quickly kill their insect host after infection. The polyhedrin protein serves to sequester, and thereby protect, hundreds of virus particles from proteolytic inactivation by the decomposing host tissue. The virus is transmitted when occlusion bodies are ingested by a new host as it feeds on a contaminated food source. The polyhedrin protein dissolves in the alkaline environment of the new host's gut and the occluded virus is released. This virus infects the gut epithelial cells and virus replication takes place. Nonoccluded virus is then produced and budded from the infected gut cells. At this point, the virus spreads throughout the tissues of its new host. Although the polyhedrin protein is essential for survival of the virus in nature, it is dispensable for virus survival and propagation in tissue culture cells.

BACULOVIRUS EXPRESSION SYSTEM

The baculovirus expression system takes advantage of several facts about polyhedrin protein: (1) that it is expressed at very high levels in infected cells, constituting more than

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY LIBRARY

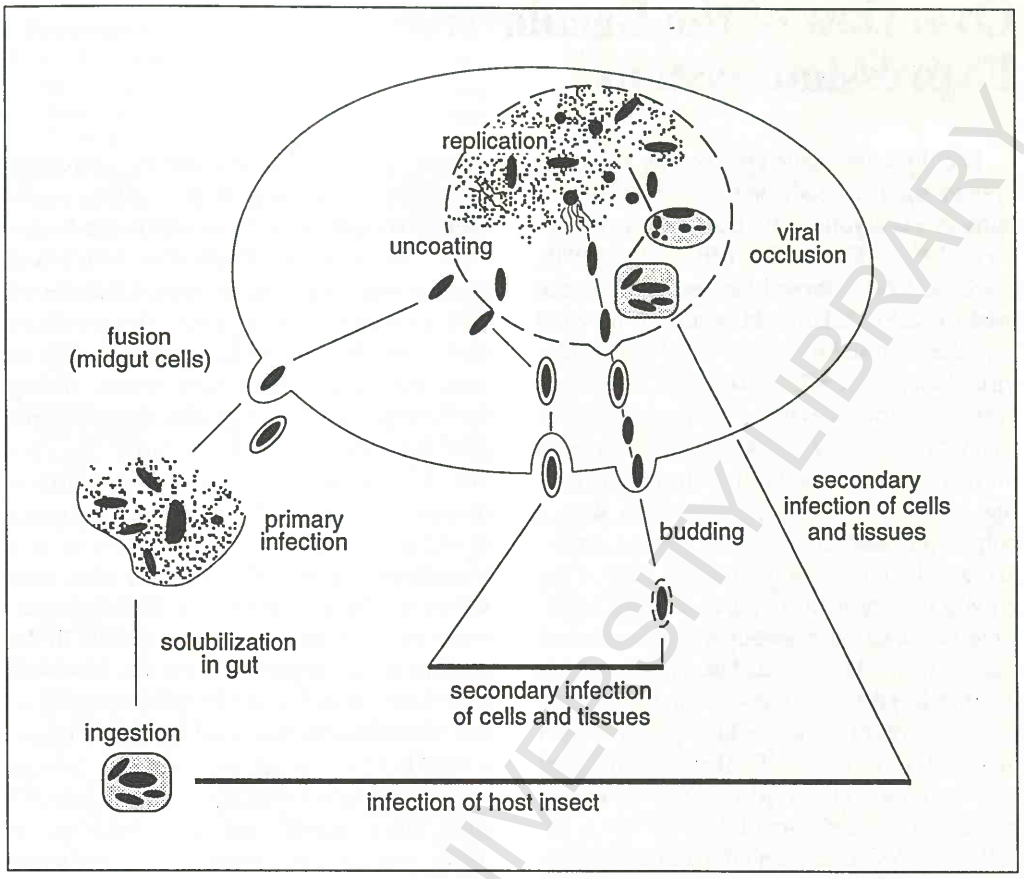


Figure 5.4.1 Baculovirus life cycle. Viruses enter cells by adsorptive endocytosis and move to the nucleus where their DNA is released. Both DNA replication and viral assembly take place in the nuclei of infected cells to generate two types of viral progeny. These include extracellular (non-occluded) virus particles and polyhedra-derived (occluded) virus particles. Extracellular virus is released from the cell by budding, starting at ~12 hr postinfection and ending ~36 hr postinfection. Polyhedra-derived virus appears later (~18 hr postinfection) and accumulates in the nuclei of infected cells ≤72 hr postinfection or until cellular lysis. Polyhedra-derived virus is embedded in proteinaceous viral occlusions, the major protein component of which is the viral polyhedrin protein. Secondary infection of cells and tissues occurs by two pathways. In the first, the extracellular virus, once budded from the site of primary infection, is free to infect neighboring cells by the pathway just described. Alternatively, polyhedra-derived virus is released from occlusion bodies after an infected food source is ingested by a new host. Reproduced from Summers and Smith (1987) with permission from the Texas Agricultural Experiment Station.

half of the total cellular protein late in the infectious cycle; (2) that it is nonessential for infection or replication of the virus, meaning that the recombinant virus does not require any helper function; and (3) that viruses lacking the polyhedrin gene have a plaque morphology which is distinct from that of viruses containing the gene. Recombinant baculoviruses are generated by replacing the polyhedrin gene with a foreign gene through homologous recombination. In this system, the distinctive plaque morphology provides a simple visual screen for identifying the recombinants.

To produce a recombinant virus that expresses the gene of interest, the gene is first cloned into a transfer vector (described below). Most baculovirus transfer vectors contain the polyhedrin promoter followed by one or more restriction enzyme recognition sites for foreign gene insertion. Once cloned into the expression vector, the gene is flanked both 5′ and 3′ by viral-specific sequences. Next, the recombinant vector is transfected along with wild-type viral DNA into insect cells. In a homologous recombination event, the foreign gene is inserted into the viral genome and the polyhedrin gene is excised. Recombinant viruses lack the polyhedrin gene and in its place contain the inserted gene, whose expression is under the control of the polyhedrin promoter.

RCU IASLIC CENTRAL UNIVERSITY LIBRARY

Homologous recombination between circular wild-type DNA and the recombinant plasmid DNA occurs at a low frequency (typically 0.2% to 5%). However, linearization of wild-type baculovirus DNA before cotransfection with plasmid DNA increases the proportion of recombinant virus to ~30% (Kitts et al., 1990). If the DNA is linearized such that an essential portion of the 1629 open reading frame (ORF) downstream from the polyhedrin gene is deleted, 85% to 99% of the viruses obtained by cotransfection with a plasmid vector that complements the deletion express the heterologous gene (Kitts and Possee, 1993). Two companies (Pharmingen and Clontech) market linear AcMNPV DNA containing such a deletion.

Once a virus stock is obtained after cotransfection, it is necessary to purify recombinant virus by plaque assay so the recombinant virus can be identified. Limiting dilution can also be used (see O'Reilly et al., 1992). One of the beauties of this expression system is a visual screen allowing recombinant viruses to be distinguished. As mentioned above, the polyhedrin protein is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded virus particles. These occlusion bodies, up to 15 μm in size, are highly refractile—i.e., they have a bright, shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. Thus, when the virus is plaqued onto *Sf*9 (*Spodoptera frugiperda*) cells, plaques can be screened for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. Recombinant viruses can also be identified by DNA hybridization and polymerase chain reaction (PCR) amplification.

POST-TRANSLATIONAL MODIFICATION OF PROTEINS IN INSECT CELLS

Because baculoviruses infect invertebrate cells, it is possible that the processing of proteins produced by them is different from the processing of proteins produced by vertebrate cells. Although this seems to be the case for some post-translational modifications, it is not the case for others. For example, two of the three post-translational modifications of the tyrosine protein kinase, pp60$^{c\text{-}src}$, that occur in higher eukaryotic cells (myristylation of the N-terminal glycine residue and phosphorylation of serine 17) also take place in insect cells. However, another modification of pp60$^{c\text{-}src}$ observed in vertebrate cells, phosphorylation of tyrosine 527, is almost undetectable in insect cells (Piwnica-Worms et al., 1990).

In addition to myristylation, palmitylation has been shown to take place in insect cells. However, it has not been determined whether all or merely a subfraction of the total recombinant protein contains these modifications. Cleavage of signal sequences, removal of hormonal prosequences, and polyprotein cleavages have also been reported, although cleavage varies in its efficiency. Internal proteolytic cleavages at arginine- or lysine-rich sequences have been reported to be highly inefficient, and alpha-amidation, although it does not occur in cell culture, has been reported in larvae and pupae (Hellers et al., 1991). In most of these cases a cell- or species-specific protease may be necessary for cleavage. Protein targeting seems conserved between insect and vertebrate cells. Thus, proteins can be secreted and localized faithfully to either the nucleus, cytoplasm, or plasma membrane. Although much remains to be learned about the nature of protein glycosylation in insect cells, proteins that are N-glycosylated in vertebrate cells will also generally be glycosylated in insect cells. However, with few exceptions the N-linked oligosaccharides in insect cell-derived glycoproteins are only high-mannose type and are not processed to complex-type oligosaccharides containing fucose, galactose, and sialic acid. O-linked glycosylations have been even less well characterized in *Sf*9 cells, but have been shown to occur. For further information on protein processing in insect cells, see Jarvis and Summers (1992) and O'Reilly et al. (1992).

LITERATURE CITED

Carbonell, L.F., Klowden, M.J., and Miller, L.K. 1985. Baculovirus-mediated expression of bacterial genes in dipteran and mammalian cells. *J. Virol.* 56:153-160.

Doerfler, W. and Bohm, P. 1986. The Molecular Biology of Baculoviruses. Springer-Verlag, New York.

Hellers, M., Gunne, H., and Steiner, H. 1991. Expression and post-translational processing of preprocecropin-A using a baculovirus vector. *Eur. J. Biochem.* 199:435-439.

Jarvis, D.L. and Summers, M.D. 1992. Baculovirus expression vectors. *In* Recombinant DNA Vaccines: Rationale and Strategies (R.E. Isaacson, ed.) pp. 265-291. Marcel Dekker, New York.

Kitts, P.A. and Possee, R.D. 1993. A method for producing recombinant baculovirus expression vectors at high frequency. *BioTechniques* 14:810-817.

BUTLER/CETRE JOINT LIBRARY

Kitts, P.A., Ayres, M.D., and Possee, R.D. 1990. Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucl. Acids Res.* 18:5667-5672.

Luckow, V.A. 1991. Cloning and expression of heterologous genes in insect cells with baculovirus vectors. *In* Recombinant DNA Technology and Applications (A. Prokop, R.K. Bajpai, and C. Ho, eds.) pp. 97-152. McGraw-Hill, New York.

Luckow, V.A. and Summers, M.D. 1988. Trends in the development of baculoviral expression vectors. *Bio/Technology* 6:47-55.

Miller, L.K. 1988. Baculoviruses as gene expression vectors. *Annu. Rev. Microbiol.* 42:177-199.

Miller, D.W., Safer, P., and Miller, L.K. 1986. An insect baculovirus host vector for high-level expression of foreign genes. *In* Genetic Engineering, Vol. 8 (J.K. Setlow and A. Hollaender, eds.) pp. 277-298. Plenum, New York.

O'Reilly, D.R., Miller, L.K., and Luckow, V.A. 1992. Baculovirus Expression Vectors. W.H. Freeman and Company, New York.

Piwnica-Worms, H., Williams, N.G., Cheng, S.H., and Roberts, T.M. 1990. Regulation of $pp60^{c\text{-}src}$ and its association with polyoma virus middle T antigen in insect cells. *J. Virol.* 64:61-68.

Summers, M.D. and Smith, G.E. 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station Bulletin No. 1555. College Station, Texas.

Key Reference

O'Reilly et al. 1992. See above.

A guide assembled to aid researchers using the baculoviral expression system, containing detailed protocols for using this system effectively.

Contributed by Cheryl Isaac Murphy
Cambridge Biotech Corporation
Worcester, Massachusetts

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Protein Expression in the Baculovirus System

Insect cell–recombinant baculovirus co-cultures offer a protein production system that complements microbial systems (*UNITS 5.1-5.3*) by providing recombinant proteins in soluble form and with most post-translational modifications. Moreover, the large size of the viral genome (130 kbp) enables cloning of large segments of DNA and consequent expression of complex protein aggregates. This unit describes methods associated with the large-scale production of recombinant proteins in the baculovirus expression system. *UNIT 5.4* gives an overview of this system. Performing the protocols in this unit requires that the desired protein has already been cloned into an appropriate vector for expression in cultured insect cells under the control of a strong baculoviral promoter; it is also assumed that both infection of cells by the recombinant virus and production of recombinant proteins have been analyzed in preliminary experiments. Detailed protocols for the construction of a recombinant baculovirus, maintenance of insect cells in culture, and analysis in small-scale expression experiments can be found in several manuals (Murphy and Piwnica-Worms, 1994a,b; Summers and Smith, 1987).

Production of proteins in large amounts requires correspondingly large quantities of starting materials, most importantly viral stocks of high titer. Basic Protocol 1 gives a method for large-scale production of viral stocks. Quality control is absolutely vital, and methods for titration of virus are provided (Support Protocols 1 and 2 describe the plaque assay and the end-point assay, respectively). Moreover, once viral stocks have been prepared and titered, the virus is tested in small-scale cultures to determine the kinetics of expression (Basic Protocol 2), which allows evaluation of various cell culture and infection conditions aimed at developing optimal levels of protein production (e.g., comparisons of different host cell lines, media, and environmental parameters). Support Protocol 4 provides instructions for preparing culture samples for protein analysis by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE; *UNIT 10.1*), and Support Protocol 5 discusses analytical methods for monitoring nutrient levels in cell culture fluids.

Once optimal process parameters are identified, the target protein can be produced on a large scale in fermentors. Basic Protocol 3 describes regular batch production in bioreactors, whereas Alternate Protocol 1 covers a variant fed-batch procedure of production in perfusion cultures. Techniques for harvesting cultures from bioreactors are provided in Support Protocol 3. These protocols produce cell pellets or culture supernatants containing the protein of interest, which must then be subjected to purification and analysis as described elsewhere in this manual (Chapters 6 and 7).

Despite being based on the use of a virus, the expression of recombinant protein with the baculovirus system is not hazardous to humans, as the virus cannot infect cells of noninsect origin. Nonetheless, handling of virus stocks should be restricted to well-defined equipment, materials, and rooms, kept separate from those where the host cell lines are maintained. This helps prevent any uncontrolled latent viral infection of the host cell line stocks.

NOTE: All protocols (except Support Protocol 4 and Support Protocol 5) should be carried out using strict sterile technique, in the absence of antibiotics.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

Copyright © 1995 by John Wiley & Sons, Inc.

BASIC PROTOCOL 1

LARGE-SCALE PRODUCTION OF VIRAL STOCK

The first phase of large-scale production of recombinant proteins is the generation of a large volume (1 to 5 liters) of highly infective recombinant virus. A culture of insect cells is prepared, grown to mid-exponential phase, and infected with recombinant virus. On day 4 or 5 postinfection, the culture is harvested by centrifugation and then stored. A small sample of the culture is reserved for titering by the plaque assay or by the end-point dilution method (see Support Protocols 1 and 2). High-titer stocks are used in the analysis of expression kinetics (see Basic Protocol 2) and in the infection of reactor cultures (see Basic Protocol 3 and Alternate Protocol 1).

Materials

Spodoptera frugiperda (*Sf*9) cells
Insect cell medium: TC100 (GIBCO/BRL) or IPL41 (GIBCO/BRL), with 5% to 10% (w/v) fetal calf serum (FCS); or Sf-900 II (GIBCO/BRL), serum free
Pure recombinant virus (at 10^7 pfu/ml)

1-liter to 10-liter capacity shaker flasks or spinner culture flasks
27°C shaker incubators, humidity controlled
Low-speed centrifuge (Sorvall RC3B or equivalent)
1-liter glass conical-bottom centrifuge bottles (Corning or Bellco), sterile

Additional reagents and equipment for monitoring cell growth by cell counting and determining cell viability with trypan blue (see Support Protocol 6), monitoring nutrient levels in culture media (see Support Protocol 5), and titration of virus (see Support Protocols 1 and 2)

1. Prepare a 500-ml culture of *Sf*9 cells in a shaker flask or spinner culture flask containing prewarmed insect cell medium (with or without serum) in a 27°C shaker incubator under controlled humidity.

   Always use a low liquid/total volume ratio (<0.4) to ensure good aeration of the culture while minimizing mechanical shear effects. Pluronic F-68 (BASF) should be included at 2 g/liter in media that lack a shear protectant. Set agitation to 50 to 100 rpm to ensure good mixing and aeration.

   To avoid uncontrolled infection, noninfected cultures should be handled separately and kept physically apart from infected cultures (use different incubators and different microbiological hoods).

   The Sf9 cell line should be maintained and regularly passaged two to three times per week in either monolayer cultures or small shaker cultures in the same medium as the one chosen for the protocol. Maintainance of the cell line is described in Murphy and Piwnica-Worms (1994a).

2. Withdraw a sample of culture and perform appropriate tests to monitor cell growth, viability, and nutrient levels to ensure an adequate growth environment.

   Growth should be monitored by cell counting using a hemocytometer and microscope. The culture should be initiated with a minimum of 2-3 × 10^5 cells/ml, at a viability higher than 98%. Healthy cells of high viability are characterized under the microscope by regular spherical shapes, smooth membranes, and refractile nuclei. Viable cells exclude trypan blue dye, whereas dead cells don't (see Support Protocol 6). Elongated cell shapes or rough cytoplasmic membranes may be an indicator of oxygen limitation (this can be relieved by increased agitation or reduced culture volume/vessel volume ratios). Nutrient levels are measured using an assortment of specific assays (see Support Protocol 5). Typically, glucose and glutamine levels should not decrease below 0.2 g/liter and 0.2 mM, respectively.

3. When the culture reaches mid-exponential phase (~ 1 to 5 × 10^6 cells/ml depending on the cell line and medium employed), infect with an aliquot of the pure recombinant

BCU IASI CENTRAL UNIVERSITY LIBRARY

virus to achieve a low multiplicity of infection (0.01 to 0.1). Transfer the spinner or shaker flask to a biological hood reserved for handling virus-infected cultures. Carefully open the flask and using a sterile pipet, add the appropriate volume of virus suspension to the culture. Mix the contents of the flask by gentle swirling, close the flask, and return it to the incubator.

For a 500-ml culture at 1×10^6 cells/ml, 5 ml of a virus suspension at 10^7 pfu/ml is appropriate.

The multiplicity of infection (MOI) is the ratio of infective virus particles per cell. The volume of virus solution (V_v, in ml) to add to a culture volume (V_c, in ml) is calculated as $V_v = MOI \times N \times V_c/T$, where N is the cell density in the culture (cells/ml) and T is the titer of the virus solution (pfu/ml).

4. Incubate culture 4 to 5 days in a 27°C shaker incubator under controlled humidity to allow virus production. Monitor cell growth once per day.

   At 4 to 5 days postinfection a large proportion of the cells (>50%) display hallmarks of infection: large cell bodies, condensed nuclei, rough membranes, and, for some cell lines, sausage shapes. At this stage, viability should be <20%.

5. Harvest the culture by centrifugation in sterile conical-bottom bottles for 30 min at $200 \times g$, 4°C, to remove cell debris. Retain the supernatant as the viral stock.

   Before discarding cell debris, inactivate all viruses in the pellet by soaking 30 min in 1 N NaOH.

6. Withdraw a small sample (e.g., 5 ml) for viral titration and store remaining stock at 4°C, protected from light, in 175-cm^2 tissue culture flasks.

   Viral stocks can be stored in the dark for several months at 4°C without significant loss of infectivity. Low-titer viral stocks ($<5 \times 10^7$ pfu/ml) may be more unstable, and titration of existing viral stocks should be performed immediately before use. Exposure to light has been shown to reduce virus infectivity during long-term storage (Jarvis et al., 1994).

   Two to three aliquots of 1 ml each can be dispensed in sterile cryovials under the biological hood (see UNIT 5.2, Support Protocol 3), tightly closed, and stored in liquid nitrogen as a back-up source for the virus. Freezing of the virus suspension, however, may lead to loss of virus titer. This loss, combined with the small volume that is frozen, would not eliminate the need to go back to step 1 of this protocol to generate a high-titer virus stock.

7. Estimate the titer of the reserved stock sample either by the plaque assay or by the end-point dilution method (see Support Protocols 1 and 2).

   Stocks of $<10^7$ pfu/ml are useless and should be discarded. $1\text{-}5 \times 10^7$ pfu/ml are considered to be low titer. The expected average titer is 10^8 pfu/ml.

DETERMINATION OF EXPRESSION KINETICS

BASIC PROTOCOL 2

Once a viral stock is made with a titer above 5×10^7 pfu/ml, the kinetics and the level of protein expression are characterized in small-scale cultures in terms of the optimal cell culture and infection parameters. The following protocol can be used to compare host cell lines, media, virus constructs, and usual process parameters (e.g., cell density, pH, and agitation). It involves growing small-volume cultures of insect cells, infecting with virus stock, and periodically withdrawing culture samples that are prepared for protein analysis (see Support Protocol 4). The results are used to guide the timing of harvesting in large-scale production systems (see Basic Protocol 3 and Alternate Protocol 1).

BCM LABS / CENTRAL UNIVERSITY LIBRARY

Materials

Cell line: *Spodoptera frugiperda* (*Sf*9, *Sf*21; ATCC, Pharminogen, Invitrogen) *or Trichoplusia ni* (*Tn*5 line; High Five cell line, Invitrogen)
Insect cell medium: TC100 (GIBCO/BRL) or IPL41 (GIBCO/BRL), serum free or with 5% to 10% (w/v) fetal calf serum (FCS); or Ex-Cell 401 or Ex-Cell 405 (JRH Biosciences) or Sf-900 II (GIBCO/BRL), serum free
Recombinant viral stock (see Basic Protocol 1)

1.5- to 3-liter shaker flasks or spinner culture flasks
27°C incubator, humidity controlled

Additional reagents and equipment for monitoring cell growth by cell counting and determining cell viability with trypan blue (see Support Protocol 6), monitoring nutrient levels in culture media (see Support Protocol 5), and preparing samples for protein analysis (see Support Protocol 4)

1. Prepare a culture of cells in a 1.5- to 3-liter shaker flask or spinner culture flask containing 0.5 to 1 liter insect cell medium (with or without serum). Place flask in a 27°C incubator and allow cells to grow. Monitor cell growth, cell viability, and nutrient levels.

   Growth should be monitored by cell counting using a hemacytometer and microscope, and cell viability is determined by the trypan blue exclusion method (see Support Protocol 6). Nutrient levels are measured using an assortment of specific assays (see Support Protocol 5).

   In serum-containing media, the cell density should not exceed 1×10^6 cells/ml at time of infection. In serum-free media, infection can be started at up to 5×10^6 cells/ml. Prepare separate test cultures for each cell culture and infection variable to be analyzed (e.g., MOI, cell line, medium, and cell density).

2. Infect culture at high MOI (5 to 10) by sterile addition of an aliquot of recombinant viral stock. Return culture to 27°C incubator and incubate under controlled humidity.

   To calculate the amount of virus stock to add, use the calculations given in Basic Protocol 1, step 3. For example, for a 500-ml culture at 10^6 cells/ml, adding 5 ml of a virus stock (at 10^8 pfu/ml) would give an MOI of 1; 50 ml of the same stock would result in an MOI of 50.

   In some cases, it may be advantageous to centrifuge the cells and suspend them in fresh medium just prior to infection. This avoids possible nutrient limitation and toxic by-product inhibition during infection.

3. Withdraw 10-ml culture samples, measure nutrient levels (see Support Protocol 5), and prepare aliquots for protein analysis (see Support Protocol 4). Monitor nutrient levels and cell viability using trypan blue on a regular basis (twice per day) until 4 to 5 days postinfection.

   Once infected, cells are not necessarily completely permeable to trypan blue, and intermediate levels of staining can be observed. Also, some cell lines still exclude trypan blue quite effectively even when fully infected. An increase in cell size is usually a better indicator of infection.

4. Determine the optimal harvest time from the kinetics of target protein expression.

   SDS-PAGE (UNIT 10.1), immunoblotting, and biological assay, if available; may be used to analyze protein levels. For some proteins, it may be beneficial to harvest before the actual peak of expression because of progressive proteolytic degradation. Such degradation may be apparent as undersized immunoreactive bands in blots.

   Because the infection has been carried well beyond the optimal harvesting time, and unless the desired protein is expressed at very high levels and is very stable, the test cultures cannot be used further. They should be neutralized for 30 min in 1 N NaOH.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

PRODUCTION IN BIOREACTORS

This protocol describes large-scale production of recombinant proteins in insect cell cultures conducted in bioreactors. Equipment preparation involves inspecting the reactor for cleanliness, calibrating sensors, and sterilizing. Culture medium is introduced through a sterilizing filter and then inoculated with exponentially growing, uninfected insect cells prepared in shaker or spinner flasks (see Basic Protocol 1, steps 1 and 2). Once cells reach half-maximal density, viral stock is introduced. The culture is terminated at the appropriate time, as determined in small-scale experiments (see Basic Protocol 2), and harvested (see Support Protocol 3). Because cultures are performed for days or weeks in the absence of antibiotics, maintaining the sterility of all operations is essential. This method may be adapted to perfusion cultures (see Alternate Protocol 1).

Materials

Insect cell medium: TC100 (GIBCO/BRL) or IPL41 (GIBCO/BRL), serum free or with 5% to 10% (w/v) fetal calf serum (FCS); or Ex-Cell 401 (JRH Biosciences) or Sf-900 II (GIBCO/BRL), serum free

Inoculum (15% to 20% bioreactor culture volume): spinner or shaker flask culture of uninfected insect cells and in full exponential growth phase

Recombinant viral stock (see Basic Protocol 1)

Bioreactor fitted with temperature, pH, agitation, air flow, and pO_2 control (see Table 5.3.1 and Fig. 5.3.1)

Millipak filter, sterile (Millipore)

Viral stock and inoculum transfer vessel fitted with transfer line (see Fig. 5.3.2)

Peristaltic pump (optional)

Additional reagents and equipment for checking sterility and monitoring cell growth and determining cell viability with trypan blue (see Support Protocol 6)

1. Prepare bioreactor by inspecting for the absence of residual biological material on the reactor vessel walls, agitator shaft baffles, or other parts; calibrating sensors for temperature, pH, pO_2 and pressure; and sterilizing accessory lines and transfer vessels by autoclaving. Fill reactor vessel with water such that all probe tips are immersed and sterilize. Empty reactor.

   If the bioreactor is autoclaved, it is important to ensure proper monitoring of the temperature inside the reactor (UNIT 5.3).

2. Connect medium addition line and add medium through a sterilizing filter. Disconnect medium addition line.

   The volume of the medium added is 75% to 80% of the final culture volume desired. This allows for cell inoculum and viral stock additions.

3. Set all parameters to the respective set points.

   Among strategies available to control dissolved oxygen levels in the culture (e.g., flow rate of gas, agitation, control through headspace), pure O_2 pulsing through the sparger is the easiest method to monitor. A permanent airflow through the reactor headspace (at 1 liter air/liter culture/min) also damps possible oscillations around the set point. Generally, a set point of 30% to 50% air saturation should be optimal.

   If possible, leave the bioreactor in an uninoculated state for 24 to 48 hr. Check sterility to detect contaminants and to avoid loss of valuable inoculum.

4. Inoculate bioreactor with inoculum consisting of a spinner or shaker flask culture of uninfected insect cells at high viability and in full exponential growth phase so as to obtain a starting density of 2 to 3×10^5 cells/ml. Pool cultures if necessary and transfer

BOJAS / CENTRAL UNIVERSITY LIBRARY

into a sterile inoculation transfer vessel under a biological hood. Connect transfer vessel as described in *UNIT 5.3*.

Use a septum port that has not been previously used for emptying or filling the reactor.

The growth characteristics of the inoculum should have exhibited predictable and reproducible growth rates during previous subculture.

5. Transfer the inoculum aseptically into the reactor either via a peristaltic pump or by pressurizing the inoculum vessel with sterile air.

   Foaming of the culture should be avoided at all times during the transfer.

6. Monitor cell growth until cell density reaches half-maximal levels, as estimated from inoculum growth curves.

7. Infect culture at high MOI (5 to 10) by aseptically adding recombinant viral stock as described for inoculum addition in step 5. Allow virus infection, virus replication and protein expression to proceed for 4 to 5 days, withdrawing samples twice daily and determining cell viability and nutrient levels (see Support Protocol 5). In addition, prepare aliquots of the sample twice daily for protein analysis (see Support Protocol 4). Terminate the process by harvesting at the optimal time point (see Basic Protocol 2 regarding determination of this time point).

   The multiplicity of infection (MOI) is the ratio of infective virus particles per cell. The volume of virus solution (V_v in ml) to add to a culture volume (V_c in ml) is calculated as $V_v = MOI \times N \times V_c/T$, where N is the cell density in the culture (cells/ml) and T is the titer of the virus solution (pfu/ml).

8. Harvest by centrifugation or tangential flow microfiltration (see Support Protocol 3).

   Depending on the target protein localization, keep the cell pellet or the supernatant fraction. Representative samples should be collected before and after any of the following downstream processes (see Support Protocol 3 and Support Protocol 4) to monitor changes in target protein stability or conformation.

ALTERNATE PROTOCOL 1

PRODUCTION IN PERFUSION CULTURES

Perfusion cultures permit high cell densities and production of greater protein amounts than simple batch methods (see Basic Protocol 3), particularly if proteins are designed for secretion. Equipment setup is as in Basic Protocol 3 except that a tank containing extra medium is set on a load cell and connected to the reactor, and a cell retention device is included along with a holding tank for spent medium. After the reactor is inoculated with insect cells, fresh medium is added to the culture at increasing rates to match the increase in cell density. Spent medium (free of cells) is withdrawn at the same rate to keep reactor volume constant. Because of the large holding tanks required, this protocol can be easily implemented at reactor scales of 1 to 5 liters but requires an industrial production environment for larger scales.

Additional Materials (also see Basic Protocol 3)

Cell retention device: spin filter, internal membrane filter, external membrane filter, or external centrifuge
Fermentor level controller (contact probe)
Medium feed and filtrate polypropylene holding tanks (10 to 15 times reactor volume; Nalgene)
Load cell or electronic balance for feed tank (Mettler or equivalent)
Peristaltic pumps (Watston-Marlow or equivalent)
Additional reagents and equipment for determining cell viability with trypan blue

BCU IASL CENTRAL UNIVERSITY LIBRARY

(see Support Protocol 6), and preparing samples for protein analysis (see Support Protocol 4)

1. Prepare the bioreactor and start the culture (see Basic Protocol 3, steps 1 to 6). Be sure to equip the bioreactor with a cell retention device and a fermentor level controller.

   Set up the level controller so that it activates addition of fresh medium to the reactor by a peristaltic pump as culture volume decreases.

   The decrease in culture volume is caused by continuous withdrawal of cell-free filtrate by another pump. The speed of this second pump determines the perfusion rate (0.5 to 4 culture volumes/day).

   The equipment used here is available from most fermentor manufacturers and can be adapted to standard batch-oriented designs (see Table 5.3.1). For a thorough review of perfusion hardware, see Tokashiki and Takamatsu (1993).

2. Prepare a large volume of medium in a feed tank. Set tank on load cell or electronic balance.

   Feed and filtrate holding tanks should be of at least 10× the reactor capacity. This explains why it is more difficult to run large reactors (more than 5 to 6 liters) in perfusion in a common laboratory environment.

3. Connect feed tank to the reactor and connect filtrate holding tank to the filtrate port of the cell retention device. Set up peristaltic pumps for medium feed and filtrate removal.

4. Monitor cell growth, and, once cell density has reached half the maximal density for a batch operation, start perfusion with insect cell medium at a rate of 0.5 reactor volume/day.

   Monitor feed rates by recording changes in the weight of the feed tank per unit time (be sure to take into account the density of the feed medium).

5. Gradually increase perfusion rate in proportion to the increase in cell density. Allow 3 to 5 days of perfusion to reach the target cell density.

   The target cell density for infection is 1 to 3×10^7 cells/ml at more than 95% viability.

   Ideally, a perfusion process should be run without infection first to determine the maximal cell density achievable with a given set of hardware, cell line, and medium. However, most reports available in the literature seem to agree with a maximal cell density of 5-6 $\times 10^7$ cells/ml.

6. Infect the culture by adding recombinant viral stock to the reactor (see Basic Protocol 3). Withdraw a 5-ml sample for analysis of protein expression. Halt perfusion for a short time after infection (1 to 2 hr) to allow virus attachment and/or penetration inside the cells.

   This sample will serve as a negative control in protein expression analysis (see Support Protocol 4).

   In this process, because the cell density is unusually high, it may be difficult to achieve a high MOI. However, it has been shown that infection at very low MOI (0.01) is still as productive as a batch operation with a high MOI.

   Perfusion is temporarily halted to prevent virus escape through the cell retention device. To avoid nutrient limitation during the arrest of perfusion, the feed rate can be increased 2- to 3-fold during the hour preceding the infection. Halting perfusion can simply be done by turning off the filtrate withdrawal pump. The rest of the cell retention hardware (filters, pump, centrifuge) should be left running without interruption to avoid any heterogeneity of the whole system.

BOKU WIEN / ZENTRALE UNIVERSITÄTS BIBLIOTHEK

7. Set perfusion rate to maximum (3 to 4 reactor volumes/day).
8. Let infection proceed for at least 4 days. Monitor cell viability and protein expression.
9. To avoid dilution of the product in the filtrate (and if the protein is secreted), remove the accumulated filtrate every 24 hr and concentrate by ultrafiltration (see Support Protocol 3). Prepare sample for analysis (see Support Protocol 4).

If expression of the secreted protein is under control of a late promoter, the accumulated filtrate for the first 24 hr postinfection usually contains little product and can be discarded.

SUPPORT PROTOCOL 1

TITRATING VIRAL STOCKS WITH THE PLAQUE ASSAY

It is important to know the titer of a viral stock when preparing new stocks or when infecting insect cells to produce recombinant proteins. Two methods are used to determine viral titers in stock preparations, namely, the plaque assay and the end-point titration assay (see Support Protocol 2). Both methods are biological assays employing cell culture and quantify only the extracellular infectious forms of baculovirus. The plaque assay involves infecting exponentially growing *Sf*9 cell monolayers with serial dilutions of viral stock (see Basic Protocol 1). After removing the viral supernatant following a period of viral adsorption, an agarose overlay is added and culture is continued for 5 to 7 days before plaques are counted in the culture treated with the highest dilution of virus that elicits an infection.

Materials

*Sf*9 cells (30 ml, 8×10^5 cells/ml)
TC100 medium (GIBCO/BRL), with and without 5% (w/v) fetal calf serum (FCS)
Test sample: viral stock or culture supernatant
3% (w/v) low gelling/melting temperature agarose, autoclaved
2× TC100 medium/10% (w/v) FCS
0.3% (w/v) neutral red, diluted 1:20 in 1× phosphate-buffered saline (PBS, *APPENDIX 2E*; filter sterilize and store in the dark)

6-well cluster dishes, sterile
Pipet tips, sterile
Tubes, sterile
Humid box

1. Seed three sterile 6-well cluster dishes by placing 1.5 ml of exponentially growing (8×10^5 cell/ml) *Sf*9 cells in TC100 medium containing 5% FCS in each well. Place dishes in incubator at 27°C.
2. After monolayer formation (after 30 min to 1 hr incubation), wash cells with TC100 medium (no serum).

Serum-free medium facilitates the adsorption of the virus to the cells.

3. Using sterile pipet tips and tubes, dilute test sample to 10^{-6}, 10^{-7}, and 10^{-8} in serum-free (TC100 medium). Add 300 µl of each dilution to each well of the host cell monolayers. Six repeats of one dilution per plate should be performed. Hold the pipet tip against the wall of the well while dispensing and carefully tilt the plate by hand to ensure an even distribution of the inoculum. Make six replicates of each dilution per plate.
4. Return plates to incubator and let the inoculum virus adsorb 2 hr at 27°C. During incubation, prepare the agarose overlay by melting autoclaved 3% low gelling/melt-

BCU IASI CONTRACT UNIVERSITY LIBRARY

ing temperature agarose in a microwave, cooling to 37°C, and mixing 1:1 with 2× TC100 medium/10% FCS at 37°C.

5. Remove the virus by vacuum pipetting.

   Again, hold the pipet tip against the wall of the well, plate tilted.

6. Rapidly cover with 2 ml agarose overlay per well, and let the agarose harden for 3 to 4 hr at 27°C. Overlay with 1 ml TC100 medium/5% FCS.

   The agarose prevents diffusion of virus released from infected cells over the course of the assay. Thus, only cells immediately adjacent to the cell of primary infection are subsequently infected by progeny virus. When enough cells in the immediate vicinity of the primary infection are lysed, a plaque (or hole in the cell monolayer) results.

7. Incubate plates 5 to 7 days in a humid box in the 27°C incubator.

8. Stain plates for 2 hr with 0.3% neutral red. Invert the plate on absorbent paper to decant excess liquid and store 5 hr in the dark at room temperature to allow plaques to become visible. Count plaques and calculate average of all wells showing clear, separate plaques. Express titers as pfu/ml original test sample.

   Staining of cells is not essential but facilitates plaque observation. Uninfected cells stain red, and areas of cell lysis remain clear due to plaque formation.

TITRATING VIRAL STOCKS WITH THE END-POINT ASSAY

SUPPORT PROTOCOL 2

Titrating virus by the end-point assay is similar to the plaque assay (see Support Protocol 1), but here a single virus can replicate and spread throughout the culture well and ultimately infect all cells present in the well, leading to the cytopathic effect. The titer is first given in $TCID_{50}$ units/ml, which can be transformed to an approximate result in pfu/ml. Infection of cells is assayed by detecting either the recombinant protein or the cytopathic effect.

Materials

*Sf*9 cells (1×10^6 cells/ml), grown in Sf-900 II medium (GIBCO/BRL)
Test sample: viral stock or culture supernatant
Chromogenic substrate for detecting recombinant protein (optional, if available; e.g., Xgal for β-galactose)

96-well microtiter plate, sterile
Multitip pipettor
Pipet tips, sterile
Humid box

1. Add 225 µl exponentially growing *Sf*9 cells (1×10^6 cells/ml, in Sf-900 II medium) to every well of a sterile 96-well microtiter plate.

   Do not allow the plates to stand.

2. Immediately add 25-µl undiluted test sample of virus to the first well.

   Make three replicate dilutions of each sample.

3. Make 10-fold serial dilutions across the plate, directly into the insect cell suspensions, using a multitip pipettor and changing pipet tips between each dilution to prevent carryover of virus on the outside of the tip to the low dilutions.

   Alternatively, dilutions can be made in sterile 1.5-ml microcentrifuge tubes.

B. C. U. „M. [illegible] IAŞI
C H [illegible]

BCU IASI CENTRAL UNIVERSITY LIBRARY

4. Transfer 10 µl of each of the 10^{-6} to 10^{-8} dilutions to a microtiter plate.

   Up to 60 replica wells per dilution can be made.

5. Incubate microtiter plates in a humid box 6 days at 27°C.

   It is important that no evaporation be allowed to occur due to the small volume in each well. Wrapping parafilm around the plate also avoids desiccation.

6. If available, add chromogenic substrate for detecting recombinant protein. Score wells such that evidence of viral infection is taken as a positive result.

   Evidence of viral infection may be simple in the case of virus coding for recombinant protein that can be easily detected (e.g., via the use of chromogenic substrates). Otherwise a positive well is scored on the basis of the presence of the cytopathic effect.

7. Count the number of positive (infected) and negative (uninfected) wells and calculate the theoretical dilution that would yield 50% positives, following the method described in Summers and Smith (1987). To convert $TCID_{50}$ into pfu, multiply by ln(2) (i.e., 0.69).

SUPPORT PROTOCOL 3

HARVESTING

Harvesting of cell cultures can be done by centrifugation or tangential flow microfiltration to separate cell bodies from culture medium. Centrifugation can be done in a batch mode by dividing the material into large centrifuge bottles and processing conventionally. Continuous feed centrifugation requires a continuous feed centrifuge and is effective for volumes of 5 to 20 liters. Larger volumes must be harvested by tangential flow microfiltration, which provides a retentate that must then be transferred to centrifuge bottles and subjected to batch centrifugation. Depending on the target protein localization, either the cell pellet or the supernatant fraction is kept for further purification. Protease inhibitors should be added in all cases and temperature should be kept at ~4°C to limit possible degradation. In practice, however, it becomes expensive to add them before the material is concentrated to a small volume.

Materials

Protease inhibitor mix (see recipe)
1-liter conical-bottom centrifuge bottles (Bellco), sterile
Low-speed centrifuge (Sorvall RC3B or equivalent), continuous feed centrifuge (Heareus Contifuge or equivalent), *or* tangential flow microfiltration system (Microgon KrosFloII or equivalent; 0.04 m^2/liter culture)

To harvest by batch centrifugation (≤5 liters)

1a. Transfer the reactor contents into 1-liter conical-bottom glass centrifugation bottles by pressurizing the headspace into the bottles.

2a. Centrifuge 30 min at 200 × *g*. Concentrate the cell-free supernatant by tangential flow ultrafiltration.

To harvest by continuous feed centrifugation (5 to 20 liters):

1b. Connect the inlet line of the centrifuge to the reactor harvest valve, then connect the filtrate outlet line from the centrifuge to a harvesting tank.

2b. Start centrifuge at 200 × *g*, 250 ml/min. Check that supernatant is clear and free of cells. Concentrate the cell-free supernatant by tangential flow ultrafiltration.

ICU IASI CENTRAL UNIVERSITY LIBRARY

To harvest by tangential flow microfiltration (>20 liters)

1c. Connect the inlet of the filtration system to the reactor harvest valve. Connect the retentate line to a return port at the bottom of the reactor.

It is important that the return port is immersed in the culture to avoid foaming. Alternatively, the reactor contents can be transferred to a harvesting tank, then filtered as described.

2c. Start the circulation pump with the filtrate line closed. Operate without filtration for 3 to 5 min, then open the filtrate line. When the reactor is almost empty, transfer the retentate to centrifuge bottles and centrifuge as described in step 2a. Concentrate the cell-free supernatant by tangential flow ultrafiltration.

The circulation rate and transmembrane pressure should be set so that the average shear rate is <2500 sec^{-1}, and transmembrane pressure should not exceed 500 mbar to obtain optimal filtration and avoid membrane clogging (Maiorella et al., 1991).

PREPARING SAMPLES FOR PROTEIN ANALYSIS

SUPPORT PROTOCOL 4

The preparation of samples for protein analysis depends on the protein that is expressed. For a cell-associated protein, a simple lysis in a detergent-containing buffer usually releases most intracellular and membrane proteins. Alternatively, direct resuspension in the SDS sample buffer also solubilizes all proteins. For a secreted protein, dilute the culture supernatant (after concentration if necessary) with an equal volume of sample buffer. Once prepared by this protocol, samples can be loaded directly onto polyacrylamide gels (*UNIT 10.1*).

Materials

Test sample
SDS sample buffer (*UNIT 10.1*) or detergent lysis buffer (see recipe)
Low-speed centrifuge

1. Pellet cells 5 min at 300 × *g*, 4°C.
2. Suspend in SDS sample buffer or lysis buffer (50 to 100 µl per 10^6 cells).

If the cell pellet is resuspended in lysis buffer, an aliquot of the lysate should then be diluted 1:1 with sample buffer.

3. Store at –20°C or –80°C until analyzed.

Samples can be loaded on SDS-PAGE gels (UNIT 10.1) and analyzed by direct detection (e.g., by Coomassie blue or silver staining; UNIT 10.5) of a band at the expected molecular weight with maximum intensity at optimal harvest time, or by immunodetection after blotting the gel on a membrane (UNIT 10.7). It is good practice to include negative controls (e.g., time 0 of infection, cultures infected with wild-type virus, or noninfected culture samples) as well as positive controls (e.g., partially purified recombinant or natural protein when available, or samples from other expression experiments).

BCU IASTE CENTRAL UNIVERSITY LIBRARY

SUPPORT PROTOCOL 5

MONITORING NUTRIENT LEVELS IN CULTURE MEDIA

Expression of recombinant proteins by cultured cells depends directly on levels of nutrients supplied by the culture medium. This protocol describes analytical techniques for glucose, lactate, glutamine, and ammonia. Monitoring nutrient levels is not an absolute requirement for successful exploitation of the baculovirus expression technology. It is useful, however, for troubleshooting in cases of problems with cell growth, virus stock amplification, and protein production. It allows quality control checks of media and cultures and points to eventual nutrient limitations (glucose, glutamine, and other amino acids) or inhibitory product accumulation (lactate, ammonia). A more exhaustive compendium of biochemical analytical techniques for cell culture can be found in Doyle et al., 1994.

The test sample of 5 to 10 ml is withdrawn from the culture (see Basic Protocols 1 and 2), the cells removed by brief centrifugation or filtration, and the culture fluid subjected to analysis.

Glucose and lactate. Glucose can be analyzed in an automated enzyme assay (Yellow Spring Instruments). Glucose oxidase, immobilized on a membrane, oxidizes glucose in the sample to Glucono-δ-lactone, producing H_2O_2. A platinum electrode, held at an anodic potential, in turn oxidizes H_2O_2 to produce electrons. The electron flow is measured and is directly proportional to the H_2O_2 and therefore also proportional to the glucose concentration. The same machine can analyze lactate on the same sample. Another similar membrane electrode with immobilized L-lactate oxidase at its surface gives a parallel amperometric signal proportional to the lactate concentration in the sample.

Glutamine. Glutamine is a critical nutrient for insect cell metabolism. Because glutamine degrades in liquid media, it is wise to check whether it becomes limiting for growth. Glutamine can be measured by high-performance liquid chromatography (HPLC) or by an enzyme assay in which it is first converted to glutamate and ammonia by glutaminase (Ozturk et al., 1989).

Ammonia. Ammonia can be measured by an enzymatic assay kit (Sigma) or by an ion-specific electrode (Orion 9512BN). For analysis with the ammonia electrode, the solution to be analyzed is first made alkaline by the addition of an equal volume of 0.2 N NaOH. Ammonia gas is generated which diffuses through the semipermeable membrane of the electrode and leads to a voltage potential difference. The assay is linear from 0.2 to 12 mM ammonia (log scale; Ozturk et al., 1989).

SUPPORT PROTOCOL 6

DETERMINATION OF CELL GROWTH AND VIABILITY

This protocol describes growth monitoring by cell counts and viability estimation by the trypan blue exclusion method. The culture sample is first stained with trypan blue, and an aliquot is applied to the chamber of a hemocytometer and examined under an inverted microscope. At least 200 cells should be counted to guarantee accuracy.

Materials

Fresh cell culture sample
0.2% (w/v) trypan blue solution (Sigma)
Ethanol

Hemacytometer
Inverted microscope (×10 objective, phase contrast)

1. Mix 100 µl of the fresh cell culture sample with 100 µl trypan blue solution.

ECU IASH CENTRAL UNIVERSITY LIBRARY

2. Clean the hemocytometer plate and coverslip with ethanol. Dry and fix the coverslip.
3. Fill the hemocytometer chamber with the stained sample.
4. Examine under the inverted microscope and count live (bright) and dead (blue) cells on at least two areas of 25 small squares. Average each count (live and dead) and multiply by 2×10^4 to obtain cell density/ml.

   If there are too many cells to count, dilute the original sample by 5- or 10-fold in PBS and start at step 1 again, taking dilution factor into account for the final result.

5. Calculate viability by the proportion of live cells in the total number (live + dead).

   Healthy, noninfected suspension cultures should give a viability above 95% at all times.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Detergent lysis buffer

1% (v/v) Nonidet P-40 (NP-40)
150 mM NaCl (*APPENDIX 2E*)
50 mM Tris·Cl, pH 8.0 (*APPENDIX 2E*)

Protease inhibitor mix

1 mM phenylmethylsulfonyl fluoride (PMSF)
5 mM iodoacetamide
0.2 mM *Nα-p*-tosyl-L-lysine chloromethyl ketone (TLCK; prepared fresh)
2 mM benzamidine·HCl

CAUTION: *Prepare stock solutions in a chemical fume hood.*

COMMENTARY

Background Information

The baculovirus expression system has become a popular technology to produce recombinant proteins. It is based on the infection of cultured insect cells by a recombinant virus vector in which foreign DNA is integrated under control of the strong viral promoter for polyhedrin protein (*polh*; *UNIT 5.4*). It is used for obtaining recombinant proteins for various purposes; such as biochemical studies, analysis of structure-function relationships, in vivo animal studies, and structural studies.

The system recognizes vertebrate protein targeting signals and thus allows expression of different classes of proteins: cytoplasmic, nuclear, membrane-spanning, and secreted. The host insect cell performs most post-translational modifications, including phosphorylation, acylation, glycosylation, precursor processing, subunit assembly, secretion, and nuclear translocation. The glycosylation patterns reported in the recombinant proteins, however, have been restricted to the high-mannose type in almost all cases. The baculovirus expression system is also extensively used for the construction of complex biological aggregates (e.g., ion channels, viral capsids, and multisubunit enzymes). It is biologically safe. Instructions for constructing recombinant baculovirus and small-scale expression experiments are available in other manuals (Murphy and Piwnica-Worms, 1994a,b; Summers and Smith, 1987).

The *polh* promoter is the most widely used promoter to date and many of the available transfer vectors use it. However, p10, another very late, very strong promoter is also being used (McCutchen et al., 1991; Stewart et al, 1991; Belyaev and Roy, 1993; Wang et al., 1993; Bonning et al., 1994; Deramoudt et al., 1994; van-Lier et al., 1994). In addition, the use of earlier promoters (e.g., basic protein promoter, also called *pCor*) is also exploited. With this promoter, expression of the recombinant protein starts at about 20 hr postinfection and peaks at 30 to 36 hr postinfection (Sridhar and Hasnain, 1993; Sridhar et al., 1993; Rankl et al., 1994; Chazenbalk and Rapoport, 1995). At these times, the host functions involved in protein processing are much less affected by viral infection.

BOULDER LIBRARY UNIVERSITY

Before large-scale production of recombinant proteins with the baculovirus system can be undertaken, it is necessary to prepare large-volume stocks of recombinant virus (see Basic Protocol 1). The stocks must have a known titer (see Support Protocols 1 and 2) and must be analyzed in small-scale cultures (see Basic Protocol 2) to determine optimal process parameters. Analysis includes preparing samples for electrophoresis (see Support Protocol 4) and monitoring nutrient levels in culture media (see Support Protocol 5). The results of analysis of the kinetics of recombinant protein expression are used to determine the optimal harvest time, which is a balance between maximal production and minimal degradation.

Once optimal process parameters (e.g., cell line, media, MOI, time of harvest) are identified, the target protein can be produced on a large scale in bioreactors (see Basic Protocol 3). The design of an animal cell bioreactor is similar to that of a microbial fermentor (Fig. 5.3.1) with modifications of the agitation (e.g., lower range of motor speeds, 20 to 200 rpm; marine impeller instead of Rushton turbine) and aeration (lower range of gas flow rates, headspace aeration) systems (for a review, see Hu and Wang, 1986). Equipment preparation is identical to that of a microbial reactor (*UNIT 5.3*), and harvesting is by batch centrifugation, continuous feed centrifugation, or tangential flow microfiltration (see Support Protocol 3). Because cell culture processes are run for days or weeks, without any antibiotics, special attention should be paid to the sterility of all operations.

Perfusion of insect cell cultures (see Alternate Protocol 1) is an effective technique to increase cell density and volumetric productivity of a reactor. It is particularly well suited to processes for the production of secreted proteins. The basis of the technique is to retain the cells inside the reactor by an internal spin filter, an external filter, or an external centrifuge (for a review, see Tokashiki and Takamatsu, 1993) while adding fresh medium and withdrawing spent medium at equal rates to maintain a constant reactor volume.

Critical Parameters and Troubleshooting

Insect cell line. Recombinant product expression levels are cell line–dependent (Hink et al., 1991), and it may therefore be advantageous to screen a number of cell lines for relative productive capacities. The robustness, ease of handling in suspension and in large-scale reactors, and nutrient requirements are also key parameters to consider when choosing a production cell line. Generally, *Sf*9 cells (*Spodoptera frugiperda* clone 9, derived from fall armyworm ovaries) are superior for production of virus, whereas Tn5 cells are superior for secreted proteins. From a practical point of view, it is possible to generate virus stocks in one cell line to be used for recombinant protein production in another.

Maintenance of cell lines in a predictable manner (i.e., same cell count and viability profiles over time) is extremely important to ensure reproducible and optimal recombinant product yields. As cell lines grow to different maximum densities under different medium conditions, they should be subcultured at different initial cell inoculum densities. Ideally, cell lines should be subcultured by 10-fold dilution with fresh medium and allowed to grow to densities that are ≤40% of the maximum achievable cell density (for *Sf*9 in TC100/5% FCS medium, maximum density is 2×10^6 cells/ml). Above this critical density, the culture deviates from the maximum specific growth rate and may exhibit inefficient utilization of nutrients and subsequent variability in maximal production.

Viral stocks. Two major rules are associated with stock virus preparation. (1) Keep the number of times a virus is serially passaged to a minimum (<10). Otherwise, mutant virus can become the predominant species (Kool et al., 1991), leading to a subsequent decrease in recombinant product formation in infected insect cell cultures (Wickham et al., 1991). (2) Cultures serially infected at a high MOI (5 to 10) are more likely to have higher proportions of defective interfering particles than those infected at low MOI. A high cell viability (>95%) is essential for good results. A low MOI (<0.1) allows some degree of postinfection cell growth, which leads to higher numbers of infected cells and the production of high-titer viral stocks. High-titer viral stocks are necessary to prevent the need to add large quantities of viral stock inoculum to large-scale cultures employed for recombinant protein expression.

Titration of virus. The relative advantages of the plaque assay and end-point assay may be dependent on the availability of an easy marker for infection (some viruses carry β-galactosidase or luciferase as a reporter gene). Both titration methods take at least 6 days to obtain an estimate with a large error range (50% to 100%). The end-point assay, however, is technically less difficult than the plaque assay,

ECU IASTATE UNIVERSITY LIBRARY

which allows many more replicates to be made. This can reduce the error range significantly (down to 15% to 20% for 100 replicates per dilution). In this case initial dilutions can be performed in sterile 1.5-ml microcentrifuge tubes. When, in contrast to β-gal expressing viruses, the recombinant virus does not express a protein that can be quickly and easily detected in the 96-well microtiter plates, the end-point method is more difficult to interpret, and it is advisable to use the plaque assay even at the sacrifice of accuracy.

The major hurdles in virus quantification are associated with the inherent nature of the bioassay, which depends on the condition and density of the cells. Cells from poorly growing cultures with low viability cannot be used. Even for healthy cultures, however, if the cell density used is too low, underestimation of the virus titer occurs. If the cell density is too high, nutrients essential for recombinant product expression may be consumed to a large degree by cell growth, which would also lead to the underestimation of virus titers.

Expression kinetics. Evaluation of expression kinetics needs to be performed for every recombinant virus. In particular, the optimal time of harvest is a parameter that is specific to each protein. It varies largely with the virus construct (specifically, the temporal class of the viral promoter used for driving gene expression), the type of protein (i.e., cytoplasmic versus secreted), and the extent of post-translational modifications.

Initially, infection at a higher MOI (1 to 5) is desirable to establish the optimal times of infection and harvest, especially if product instability is an issue. The MOI cannot directly increase the productive capacity of individual insect cells, but a low MOI (0.01 to 0.0001) can indirectly increase the final volumetric yield due to increased postinfection cell growth. This may be useful for the large-scale expression of stable recombinant proteins and when virus stocks are dilute. A low MOI may lead to variability of product expression due to the inherent inaccuracy of the methods employed for virus stock quantification. Slight underestimation of the true low MOI leads to excessive cell growth and increases the risk of substrate limitation. Conversely, slight overestimation of the true low MOI leads to suboptimal postinfection cell growth, and therefore decreased predicted volumetric yield.

Optimal, less variable product yields can be achieved without instability concerns by optimizing the cell density at the time of infection, employing a high MOI. The cell density at the time of infection can significantly affect production yields of proteins as well as virus. For highest yields of recombinant proteins, it is necessary to infect cultures at the highest possible cell density that supports infection and maximum specific productivites. This density usually occurs immediately prior to the mid-exponential growth phase and is dependent on the medium and cell line employed.

Maintenance of separate virus and cell banks. Physical separation of all virus-associated material and equipment from cell-associated material and equipment is essential. This includes incubators, fume hoods and associated apparatus, bioreactors, probes, cell culture reagents, and media. Accidental contamination of stock cell cultures with virus is potentially disastrous and difficult to detect. Small quantities of contaminating virus may involve a latent infection that is continuously diluted during normal subculture. If a virally contaminated culture is left to overgrow, it may be difficult to detect the infection because the culture may stop growing due to exhaustion of nutrients before a full-scale viral infection can be detected.

Production in bioreactors. Expression decreases in batch cultures infected beyond the mid-exponential phase of cell growth. Dissolved oxygen (pO_2) is the major additional parameter that can be easily controlled with bioreactors compared to spinner flasks. As a general rule, a set point of 30% to 50% dissolved oxygen tension (of oxygen contained in air) should be optimal and has been shown to be fairly independent of the target protein. Because the culture is sparged either continuously or at intervals with gas, the addition of a shear protectant such as pluronic F-68 is required to avoid mechanical damage to the cells (Murhammer and Goochee, 1988).

Basic Protocol 3 describes a procedure for stirred tank reactors. Minor modifications, in particular for pO_2 control strategy, are necessary for airlift type reactors. Rice et al. (1993) compared the two reactor types and showed that they were equivalent in terms of protein expression levels. Finally, viability should be monitored closely in large-scale systems, even when the optimal time of harvest has already been established at a small scale, to avoid excessive contamination of the supernatant with cellular proteins and to facilitate subsequent downstream processing.

Media. Recombinant product expression is nutrient dependent and therefore medium de-

PSU VAN PELT LIBRARY

pendent. Serum-free medium generally supports higher expression levels due to the incorporation of greatly increased quantities of nutrients, including carbohydrates and amino acids. For instance, Sf-900 II contains five times the amount of glucose and twice the amount of glutamine as serum-containing IPL-41 medium and is able to sustain higher maximum cell densities and recombinant product yields. Absence of serum is not directly responsible for this difference, however, as serum-supplemented IPL-41 medium is more enriched and sustains higher cell growth and protein production than serum-supplemented TC100 medium. If not included in the commercial medium, the addition of pluronic F-68 (≤4 g/liter) is essential to protect the cells against the shear stress imposed by a combined sparging and mechanical agitation in bioreactors. Fetal calf serum is available from many vendors. Different lots of serum from a number of suppliers should be obtained and tested. The lot that promotes the best growth rate and cell viability should be purchased in bulk.

Antibiotics. The routine use of antibiotics to allow contaminant-free cell culture is not recommended for two reasons. Antibiotics may affect the growth and production potential of the culture. Additionally they may lead to latent bacterial infection in the stock cell cultures with possible resistance that cannot be detected until extended cell growth or during valuable production runs.

Harvesting. Establishment of the degree of sensitivity of the target protein to degradation due to cell separation and concentration protocols should be established at small-scale prior to large-scale harvest, as it is considered to be protein specific. In the case of a secreted product, concentration by ultrafiltration should directly follow cell separation. For products that are particularly sensitive to proteolysis, we find it useful to add protease inhibitors (1 mM benzamidine, 1 mM PMSF, 5 mg/liter leupeptin, 5 mg/liter SBTI, and 5 mg/liter pepstatin A) immediately after cell separation. Degradation of cytoplasmic and secreted proteins can happen very quickly at late times of infection and make the protein difficult to purify (McGlynn et al., 1992). For a secreted protein, perfusion alleviates this problem by reducing the residence time of the protein in the bioreactor (Jäger et al., 1992).

Tangential flow microfiltration can be performed by attaching the filter system directly to the reactor (see Support Protocol 3) or, alternatively, by transferring the contents of the reactor to a harvest tank and filtering the tank contents. The filter system should have a minimum of 0.04 m^2 per liter of reactor volume so as to allow complete separation in <1 hr. Microfiltration should avoid cell breakage and release of cellular proteases (see Maiorella et al., 1991). This can be difficult considering the fragility of infected cells. It should be carried out quickly (<1 hr for large volumes).

Anticipated Results

Generally, reported levels of foreign gene expression range from about 1 to 500 mg recombinant product per 2×10^9 infected cells (Luckow and Summers, 1988; Miller, 1988). Reported expression levels of 25% to 70% total cellular protein produced in infected cultures often represent viral structural proteins such as BlueTongue virus proteins (Roy, 1990) or rabies nucleoprotein (Prehaud et al., 1990). It should be emphasized that these are examples of the potential levels, and 10 to 100 mg per 10^9 infected cells is the usual expression level range for most heterologous proteins. Comparison with different eukaryotic expression systems such as those employing mammalian cell lines indicates higher production by the baculovirus expression vector system.

Time Considerations

The baculovirus life cycle dictates that 4 to 5 days after cell infection be allowed for viral reproduction. Moreover, up to 1.5 to 2 weeks may be required to achieve target cell densities of host cells before infection.

Viral stock production takes ~4 or 5 days postinfection, and, although titration experiments can be set up in 1 day, it takes at least 6 days for titers to be determined. The time required for analysis of expression in small-scale cultures, which depends on how quickly target cell densities are achieved, ranges from 7 to 10 days. Preparing samples for electrophoresis can be accomplished in ~30 min, but SDS-PAGE or other analysis procedures require 1 to 3 days. Measuring nutrient levels in culture media can range from as short as ~15 min for simple enzyme assays employing kits or electrodes to as long as 1 day if HPLC is required.

Production in bioreactors entails similar time considerations. Preparing a reactor and loading it with medium can take half a day, and it is best to observe the system for 1 or 2 days prior to inoculation to ensure sterility. Perfusion cultures can theoretically be maintained indefinitely. The time required for harvesting batch cultures depends on the volume to be

ECU IASTATE UNIVERSITY LIBRARY

processed, ranging from 30 min for batch centrifugation to 45 min for continuous feed centrifugation to 1 hr for tangential flow microfiltration. The last procedure is followed by batch centrifugation, leading to a total time of 1.5 hr.

Literature Cited

Belyaev, A.S. and Roy, P. 1993. Development of baculo virus triple and quadruple expression vectors: Co-expression of three or four bluetongue virus proteins and the synthesis of bluetongue virus-like particles in insect cells. *Nucl. Acids Res.* 21:1219-1223.

Bonning, B.C., Roelvink, P.W., Vlak, J.M., Possee, R.D., and Hammock, B.D. 1994. Superior expression of juvenile-hormone-esterase and beta-galactosidase from the basic protein promoter of *Autographa californica* nuclear-polyhedrosis virus compared to the p10 protein and polyhedrin promoters *J. Gen. Virol.* 73:1551-1556.

Chazenbalk, G.D. and Rapoport, B. 1995. Expression of the extracellular domain of the thyrotropin receptor in the baculovirus system using a promoter active earlier than the polyhedrin promoter. Implications for the expression of functional highly glycosylated proteins. *J. Biol. Chem.* 270:1539-1543.

Deramoudt, F.X., Monnet, S., Rabaud, J.N., Quiot, J.M., Cerutti, M., Devauchelle, G., and Kaczorek, M. 1994. Production of a recombinant protein in a high density insect cell cytoflow reactor. *In* Animal Cell Technology: Products of Today, Prospects for Tomorrow. (R.E. Spier, J.B. Griffiths, and W. Berthold, ed.) pp 222-226. Butterworth.

Doyle, A., Griffiths, J.B., and Newell, D.G. 1994. Biochemistry of cells in culture. *In* Cell & Tissue Culture: Laboratory Procedures. John Wiley & Sons, Chichester, UK.

Hink, W.F., Thomsen, D.R., Davidson, D.J., Meyer, A.L., and Castellino, F.J. 1991. Expression of three recombinant proteins using baculovirus vectors in 23 insect cell lines. *Biotechnol. Prog.* 7:9-14.

Hu, W.S. and Wang, D.I.C. 1986. Mammalian cell culture technology: A review from an engineering perspective. *In* Mammalian Cell Technology (W.G. Thilly, ed.) pp. 167-199. Butterworths, London.

Jäger V., Grabenhorst, E., Kobold, A., Deutschmann, S.M., and Conradt, H.S. 1992. High density perfusion culture of insect cells for the production of recombinant glycoproteins. *In* Baculovirus and Recombinant Protein Production Processes (J.M. Vlak, E.J. Schlager, and A.R. Bernard, eds.) pp. 274-284. Roche Editions, Basel.

Jarvis, D.L. and Aurelio, G., Jr. 1994. Long-term stability of baculoviruses stored under various conditions. *BioTechniques* 16:508-513.

Kool, M., Voncken, J.W., van Lier, F.L.J., Tramper, J., and Vlak, J.M. 1991. Detection and analysis of *Autographa californica* nuclear polyhedrosis virus mutants with defective interfering properties. *Virology* 183:739-746.

Luckow, V.A. and Summers, M.D. 1988. Trends in the development of baculovirus expression vectors. *Bio/Technology* 6:47-55.

Maiorella, B., Dorin, G., Carion, A., and Harano, D. 1991. Crossflow microfiltration of animal cells. *Biotechnol. Bioeng.* 37:121-126.

McCutchen, B.F., Choudary, P.V., Crenshaw, R., Maddox D., Hammock, B.D., and Maeda S. 1991. Development of a recombinant baculo virus expressing an insect-selective neurotoxin: Potential for pest control. *Bio/Technology* 9:848-852.

McGlynn, E., Becker, M., Mett, H., Reutener, S., Cozens, R., and Lydon, N.B. 1992. Large-scale purification and characterisation of a recombinant epidermal growth-factor receptor protein-tyrosine kinase. Modulation of activity by multiple factors. *Eur. J. Biochem.* 207:265-275.

Miller, L.K. 1988. Baculoviruses as gene expression vectors. *Annu. Rev. Microbiol.* 42:177-199.

Murhammer, D.W. and Goochee, C.F. 1988. Scale-up of insect cell cultures: Protective effects of Pluronic F68. *Bio/Technology* 6:1411-1418.

Murphy, C.I. and Piwnica-Worms, H. 1994a. Preparation of insect cell cultures and baculovirus stocks. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.10.1-16.10.8. John Wiley & Sons, New York.

Murphy, C.I. and Piwnica-Worms, H. 1994b. Generation of recombinant baculoviruses and analysis of recombinant protein expression. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.11.1-16.11.19. John Wiley & Sons, New York.

Ozturk, S.S., Meyerhoff, M.E., and Palsson, B.O. 1989. Measurement of ammonia and glutamine in cell culture media by gas sensing electrodes. *Biotechnol. Tech.* 3:217-222.

Prehaud, C. Harris, R.D., Fulop, V., Koh, C.L., Wong, J., Flamand, A., and Bishop, D.H. 1990. Expression, characterization, and purification of a phosphorylated rabies nucleoprotein synthesized in insect cells by baculovirus vectors. *Virology* 178:486-497.

Rankl, N.B., Rice, J.W., Gurganus, T.M., Barbee, J.L., and Burns, D.J. 1994. The production of an active protein kinase C-delta in insect cells is greatly enhanced by the use of the basic protein promoter. *Protein Expression Purif.* 5:346-356.

Rice, J.W., Rankl, N.B., Gurganus, T.M., Marr, C.M., Barna, J.B., Waters, M.M., and Bruns, D.J. 1993. A comparison of large-scale *Sf*9 insect cell growth and protein production: Stirred vessel vs. airlift. *BioTechniques* 15:1052-1059.

BOSTON UNIVERSITY LIBRARY

Roy, P. 1990. Use of baculovirus expression vectors: Development of diagnostic reagents, vaccines and morphological counterparts of bluetongue virus. *FEMS Microbiol. Immunol.* 2:223-234.

Sridhar, P. and Hasnain, S.E. 1993. Differential secretion and glycosylation of recombinant human chorionic gonadotropin (beta-HCG) synthesized using different promoters in the baculo virus expression vector system. *Gene* 131:261-264.

Sridhar, P., Panda, A.K., Pal, R., Talwar, G.P., and Hasnain, S.E. 1993. Temporal nature of the promoter and not relative strength determines the expression of an extensively processed protein in a baculovirus system. *FEBS Lett.* 315:282-286.

Stewart, L.M.D., Hirst, M., Lopez-Ferber, M., Merryweather, A.T., Cayley, P.J., and Possee, R.D. 1991. Construction of an improved baculo virus insecticide containing an insect-specific toxin gene. *Nature* 352:85-88.

Summers, M.D. and Smith, G.E. 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experimental Station Bulletin No. 1555. College Station, Texas.

Tokashiki, M. and Takamatsu, H. 1993. Perfusion culture apparatus for suspended mammalian cells. *Cytotechnology* 13:149-159.

van-Lier, F.L.J., van-Duijnhoven, G.C.F., de-Vaan, M.M.J.A.C.M., Vlak, J.M., and Tramper, J. 1994. Continuous beta-galactosidase production in insect cells with a p10 gene based baculo virus vector in a two-stage bioreactor system. *Biotechnol. Prog.* 10:60-64.

Wang, M.Y., Kwong, S., and Bentley, W.E. 1993. Effects of oxygen/glucose/glutamine feeding on insect cell baculo virus protein expression: A study on epoxide-hydrolase production. *Biotechnol. Prog.* 9:355-361.

Wickham, T.J., Davis, T., Granados, R.R., Hammer, D.A., Shuler, M.L., and Wood, H.A. 1991. Baculovirus defective interfering particles are responsible for variations in recombinant protein production as a function of multiplicity of infection. *Biotechnol. Lett.* 13:483-488.

Key References

Summers and Smith, 1987. See above.

O'Reilly, D., Miller, L.K., and Luckow, V.A. 1994. Baculovirus expression vectors: A laboratory manual. Oxford University Press, New York.

King, L.A. and Possee, R.D. 1992. The baculovirus expression system. A laboratory guide. Chapman & Hall, London.

These manuals give excellent background information on the insect cell and baculovirus biology, expression vectors and strategies for optimization. Includes very practical aspects of many elementary but specialized procedures.

Contributed by Alain Bernard,
Mark Payton, and Kathryn R. Radford
Glaxo Institute for Molecular Biology
Geneva, Switzerland

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of Protein Expression in *Saccharomyces cerevisiae*

For more than a decade, the yeast *Saccharomyces cerevisiae* has been extensively utilized for the production of foreign proteins. Numerous characteristics of this system account for its popularity, among them (1) the existence of well-developed tools for manipulation of yeast DNA, including expression vectors and host strains; (2) a vast knowledge base regarding the genetics and biochemistry of the organism; (3) a eukaryotic secretory pathway; (4) eukaryotic post-translational modification pathways, such as those for N-linked and O-linked glycosylation; and (5) an extensive track record as a safe organism.

Foreign gene expression in *S. cerevisiae* has been reported for a wide variety of proteins derived from fungal and mammalian species, and many laboratories have contributed to a substantial base of vectors and host strains. Romanos et al. (1992) provide an extensive overview of strategies and progress toward the expression of foreign genes in *S. cerevisiae* as well as other yeasts. However, although *S. cerevisiae* has a well-developed eukaryotic secretory pathway, it is not the most efficient yeast for high-level export of proteins to the extracellular medium. There have been notable successes with respect to secretion of some foreign proteins in *S. cerevisiae* (Innis et al., 1985; Van Arsdell et al., 1987; Penttila et al., 1988; Hitzeman et al., 1990), but often transport to the extracellular medium is inefficient, especially for complex mammalian proteins of high molecular weight. Nevertheless, the application of advanced genetic and molecular biology tools should bring continued improvements in secretory capabilities of engineered *S. cerevisiae*.

As with all heterologous expression systems, the production of foreign proteins in *S. cerevisiae* poses various challenges, some of which have not yet been resolved. In general, the goals for expression of foreign proteins in *S. cerevisiae* are (1) attainment of desired protein yield; (2) production of protein with appropriate post-translational modifications, desired conformation, and biological function; and (3) secretion to the extracellular medium. Although these criteria do not apply to all expression projects, many proteins intended for human therapeutic use are normally secreted, and complex post-translational processing is required for biological function. It is especially for these cases that further engineering of *S. cerevisiae* vectors and host strains is required to improve both the product yield and product quality—particularly with regard to post-translational glycosylation, for which there are several important differences between *S. cerevisiae* and mammalian cells (Ballou et al., 1980; Kukuruzinska et al., 1987).

Table 5.6.1 Representative Vectors for Foreign Gene Expression in *S. cerevisiae*[a]

Vector	Promoter	Signal	Selectable marker	Replication platform	Reference
pTK46	*PGK1*	*MF*-α1	*URA3*	Integrated (TY-δ)	Sakai et al. (1991)
pTE432	*CUP1*	—	*TRP1*	2μm	Etcheverry (1990)
pMA230	*PGK1*	—	*LEU2*	2μm	Tuite et al. (1982)
	PHO5/GPD1	—	*LEU2*	2μm	Hinnen et al. (1989)
pMIRY 2	*GPD1*	—	*LEU2*-d	Integrated (rDNA)	Lopes et al. (1989)
pYBDT-10	*PHO5*	tPA (human)	*TRP1*	2μm	Lemontt et al. (1985)
	PGK1	levanase (*B. subtilis*)	LEU2	2μm	Wanker et al. (1995)
pYEULS1	*PGK1*	killer toxin (*K. lactis*)	*LEU2*-d and URA3	2μm	Castelli et al. (1994)

[a]Abbreviations: *CUP1*, copper-binding metallothionein; *GPD1*, glyceraldehyde-3-phosphate dehydrogenase; *MF*-α1, α-factor; *PGK1*, 2-phosphoglycerate kinase; *PHO5*, acid phosphatase; rDNA, ribosomal DNA; tPA, tissue plasminogen activator; Ty-δ, δ sequences of Ty transposable elements.

Copyright © 1995 by John Wiley & Sons, Inc.

ECU IASJ / CENTRAL UNIVERSITY LIBRARY

VECTORS AND HOST STRAINS

Through the efforts of a diverse and talented community, there are now many vectors and host strains available to direct gene expression in *S. cerevisiae* (Romanos et al., 1992). A variety of choices is available with respect to specific elements used to direct expression and secretion—e.g., the promoter to direct expression, the signal sequence for secretion, the expression cassette copy number and mechanism for replication, and the selectable marker to establish and/or maintain transformants. Some representative vectors that have been successfully used to direct foreign gene expression in *S. cerevisiae* are listed in Table 5.6.1. These examples illustrate the diversity of choices available to effect production of foreign proteins in this yeast.

Promoters

Many different promoters have been used to successfully direct expression of foreign genes in *S. cerevisiae*. In some cases, highly efficient promoters from the glycolytic pathway (glyceraldehyde-3-phosphate dehydrogenase, alcohol dehydrogenase, and phosphoglycerate kinase) have been used (Shuster, 1989). Because these promoters are among the strongest in *S. cerevisiae*, they are reasonable choices to effect high-level expression of foreign genes. However, these promoters should be used with care because expression is up-regulated by glucose, a common carbon source in yeast media.

Many tightly regulated, relatively strong yeast promoters are available to direct expression in *S. cerevisiae*, including those from *GAL1*, *PHO5*, and *CUP1*. In these cases, regulated expression can be achieved through simple manipulation of the growth medium by addition of metabolites or ions (Kramer et al., 1984)—galactose (*GAL1*), phosphate (*PHO5*), and copper (*CUP1*)—in a manner that will allow the transition from shaker flasks to large-scale fermentors. Through the use of these promoters, it is relatively simple to regulate the timing of expression, such as toward the end of a large-scale fermentation. In employing these and other expression systems, it is especially important to carefully consider the choice of host strain. For example, many host strains carry the *gal2* mutation, making the cells refractive to galactose import. Also, host strains vary considerably in their resistance to copper ions.

In addition to the naturally occurring regulated promoters, investigators have creatively utilized host strains carrying genetic alterations to achieve regulated expression of strong yeast promoters. For example, the *MF*-α1 (α-factor) promoter is constitutively expressed in wild-type cells of the α mating type. Brake et al. (1984) employed a *MAT*α strain carrying a mutation in the *sir3* gene (which regulates mating type–specific gene expression) to achieve temperature-regulated expression of epidermal growth factor driven by the *MF*-α1 promoter.

Vector Maintenance and Copy Number

In general, assuming the criterion of regulated expression can be achieved, high copy number of the expression cassette is considered desirable. High copy number can sometimes enhance expression levels and can result in greater stability of transformed yeast strains. To effect autonomous regulation of the expression cassette, many investigators choose the 2μm circle replication system. Copy number of vectors based on this system can vary depending on choice of selectable marker (see Selectable Markers) and range from ~10 to 200 copies per cell. An alternative approach to achieve both high copy number and stable replication is to integrate the expression cassette at repeated sites within the yeast genome. Lopes et al. (1989) developed an integrating plasmid, pMIRY2, that carries a portion of the rDNA of *S. cerevisiae* and can integrate at high copy number (greater than 100 copies) within this repetitive genomic region. Sakai et al. (1991) employed a strategy that led to integration of multiple copies of expression cassettes between δ sequences of Ty transposable elements on yeast chromosomes.

Transcription Terminators

Terminators from many yeast genes have been employed to efficiently terminate expression of heterologous mRNAs. It is difficult to discern clear differences in expression efficiencies based on the particular terminator employed, but it is important to note that inclusion of a transcription terminator in the expression cassette is desirable for assuring maximal effectiveness for the cassette.

Selectable Markers

An important consideration in developing expression vectors is the choice of selectable marker. This component can influence the ease of yeast transformation, copy number and stability of the expression cassette, and adaptability of the transformed strains for large-scale

UCLA BIOMEDICAL LIBRARY

fermentation. The most widely used selectable markers are yeast genes that complement auxotrophic mutations in host strains. These include, for example, the *LEU2* and *URA3* genes used in strains that are leu2 or ura3, respectively. Host strains have been assembled that carry very stable forms of these mutant loci, including strains that carry two lesions within the *LEU2* gene (*leu2-3* and *leu2-112*). In addition to maintenance based on auxotrophic complementation, positive selection is also available based on resistance to drugs such as G418 (Webster and Dickson, 1983) or to ions such as copper (Henderson et al. 1985). Positive selection is particularly useful for transformation of strains lacking auxotrophic mutations, especially industrial yeasts.

Selectable markers should be chosen carefully because the marker may have unexpected effects, e.g., on expression cassette copy number and ease of transformation of host strains. For example, two forms of the *LEU2* selectable marker are available, one of which (*LEU2*-d) carries only a portion of the *LEU2* promoter (Beggs, 1978; Erhart and Hollenberg, 1983). As a result, vectors carrying this marker are maintained at very high copy number, but transformation efficiencies are much reduced.

CHALLENGES TO THE EXPRESSION OF FOREIGN PROTEINS

Although the differences among expression vectors can be very subtle, effects on expression of foreign proteins can be quite profound. To illustrate this, results are presented from a series of studies on the expression of a human glycoprotein, antithrombin III (At-III; Strausberg and Strausberg, 1991). This example was chosen because it amply illustrates many of the challenges that can be encountered in attempting to attain high-level expression, post-translational modification, and secretion to the extracellular medium.

At-III is a human plasma protein with three disulfide bonds, three sites for N-linked glycosylation, and a nonglycosylated molecular weight of 43,000. A number of important goals were set for expression of this protein in *S. cerevisiae*, including (1) expression levels of at least several hundred milligrams per liter of culture; (2) proper signal processing to produce an amino acid sequence identical to that of the protein expressed in humans; (3) proper conformation, including the formation of the three disulfide bonds; (4) post-translational amino acid modification, especially N-linked glycosylation of three asparagine residues; and (5) secretion to the extracellular medium.

The potential of using *S. cerevisiae* for the production of glycosylated, biologically active At-III was initially tested using an expression vector carrying sequences known to efficiently direct secretion. A vector carrying the promoter and signal sequence from the well-characterized *MF*-α1 gene was chosen because this gene encodes an efficiently secreted yeast peptide, α-factor (see Fig. 5.6.1). In this case, secretion is directed by both pre and pro sequences; proteolytic processing of these pre and pro sequences is well understood both genetically and biochemically. Other important yeast elements included in the initial expression vectors for At-III production were a replication origin from the yeast 2μm circle (chosen because vectors with this origin replicate in a stable manner and can be maintained at medium or high copy number), a transcription terminator from the highly expressed glyceraldehyde-3-phosphate dehydrogenase (*GPD1*) gene for efficient termination, and the yeast *LEU2*-d gene as a selectable marker. (*LEU2*-d is a truncated form of the *LEU2* gene lacking a portion of the promoter region. An order-of-magnitude increase in vector copy number is associated with use of this selectable marker, compared to that of the full-length *LEU2* gene in *leu2* host strains.)

Yeast strains of *MAT***a** and *MAT*α mating types were transformed with this vector and several very striking results were obtained. First, it was clear from immunoblot analysis and biological activity assays that yeast can produce glycosylated At-III with biological activity. The At-III protein detected through the immunoblot appeared as a very heterogeneous band, indicating the addition of N-linked core glycosylation in the endoplasmic reticulum and outer chains of variable length in the Golgi complex. This pattern is very typical of yeast glycosylated proteins such as invertase (Kukuruzinska et al., 1987). Second, examination of the At-III expressed in yeast and treated with endoglycosidase H to remove carbohydrate (*UNIT 12.4*) revealed that the α-factor prepro sequence was retained in a substantial portion of the protein population. Third, overall expression was estimated at less than one percent of total cell protein. Interestingly, transformation with the high-copy-number vector (YpGX235) was only successful in cells of the *MAT*α mating type, where activity of the *MF*-α1 promoter was expected to be very low. This result suggested that high-level expression of At-III, ex-

PROTEIN SCIENCE LIBRARY

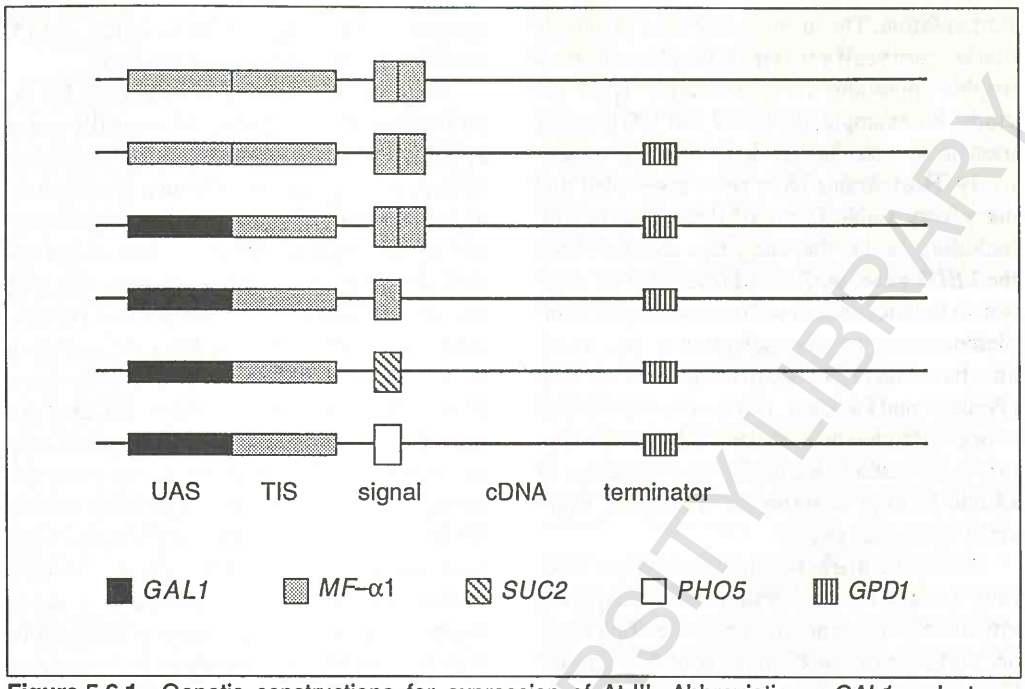


Figure 5.6.1 Genetic constructions for expression of At-III. Abbreviations: *GAL1*, galactose; *GDP1*, glyceraldehyde-3-phosphate dehydrogenase; *MF*-α1, α-factor; *PHO5*, acid phosphatase; *SUC2*, invertase; TIS, transcription initiation site; UAS, upstream activation site.

pected in *MAT*α cells, might be toxic to *S. cerevisiae*. Even at this low level of expression, most of the At-III remained cell-associated and segregated with insoluble cellular material. This result was consistent with retention of the α-factor prepro sequence.

Attainment of Desired Production Yield

To achieve regulated, high-level expression, a hybrid promoter was assembled from the yeast *GAL1* and *MF*-α1 genes. This promoter design was chosen because regulation of yeast *GAL1* is well understood and permits a variety of regulation schemes suitable for large-scale fermentation. The hybrid promoter consisted of the upstream activation site (UAS) from *GAL1* positioned upstream from the *MF*-α1 transcription initiation site. Other important elements in the vector include a combination of the 2μm origin of replication and *LEU2*-d (for high copy number) and a *GPD1* transcription terminator.

Initial expression results in yeast cells transformed with this vector were only somewhat encouraging. Although regulated expression of At-III protein was observed, the actual production levels were still quite low, possibly due to the limited availability of regulatory proteins associated with expression from the *GAL1* promoter. In this system, inducible expression is dependent upon a *trans*-acting regulatory protein, GAL4, that interacts with the UAS (Johnston, 1987). In the absence of galactose, GAL4 complexes with another protein, GAL80, and is unavailable for interaction with the UAS. Upon induction with galactose, GAL4 protein dissociates from GAL80 and binds to the UAS, thereby stimulating protein induction. It was known that the GAL4 protein is available in very limited quantities in wild-type yeast cells, so it seemed reasonable that most of the copies of the hybrid *GAL1/MF*-α1 promoter might remain dormant. Therefore, an additional modification was made so that the *GAL4* gene was positioned within the high-copy-number expression vector. The results were again only moderately encouraging because, although expression was increased several fold, the overall expression levels were no more than a few percent of total soluble protein. Immunoblot analysis revealed that ~90% of the At-III remained associated within the insoluble cell fraction.

A very striking impact on expression of At-III resulted from the use of different signal peptide coding sequences upstream from the start of mature At-III. The main goal in assessing the use of different sequences was to assess the impact on secretion, but the most noteworthy result was observed with respect to effect on expression level. Vectors were assembled

CU IAS CENTRAL UNIVERSITY LIBRARY

Table 5.6.2 Effect of Signal Sequence on Production of Soluble At-III Protein

Vector[a]	Signal	Promoter	Soluble At-III (mg/liter) produced[b]
YpGX243GAL4	*MF*-α1 prepro	*GAL1/MF*-α1	2.7
YpGX254GAL4	*SUC2*	*GAL1/MF*-α1	3.6
YpGX262GAL4	*PHO5*	*GAL1/MF*-α1	121.0

[a]Yeast strain D8 (*MATα leu2-3 leu2-112*) was transformed with each vector, and cells were grown and induced as described in Strausberg and Strausberg (1991).

[b]Determined by radioimmunoassay.

with signal sequences derived from various secreted proteins including invertase (SUC2), α-factor (MF-α1), and acid phosphatase (PHO5; see Fig. 5.6.1). For extensive discussions of signal sequences used to direct secretion in yeast see Kingsman et al. (1985) and Romanos et al. (1992). Each of the signal peptide coding sequences was assembled through insertional directed mutagenesis so that no extraneous codons were present between the signal peptide and mature protein coding sequences. For this experiment, the PHO5 signal peptide was designed predominantly from preferred yeast codons, the codons most often utilized in highly expressed yeast genes. Expression levels were monitored by radioimmunoassay and immunoblotting.

These studies were very enlightening with regard to parameters that can influence expression levels for foreign proteins in *S. cerevisiae*. First, upon induction, expression levels of soluble At-III (most of the At-III is insoluble) differed by more than an order of magnitude among the different vectors (Table 5.6.2). This result was most striking in the case of YpGX262GAL4, in which secretion is directed by the PHO5 signal. In this case, At-III protein accumulated to ~20 percent of total cell protein and microscopic examination of the yeast cells revealed particulate bodies reminiscent of the inclusion bodies associated with high-level protein expression in *Escherichia coli*. Although the mechanism for the very high level of At-III expression associated with this vector was not determined, a reasonable interpretation is that the use of preferred codons in the PHO5 signal peptide coding sequence enhanced translation initiation.

Production of Protein with Appropriate Post-translational Modifications, Conformation, and Function

One of the most difficult issues concerning production in yeast of glycoproteins intended for human therapy is the difference in glycosylation patterns for both N- and O-linked glycosylations. Although N-linked glycosylation in the endoplasmic reticulum in yeast is quite similar to that in mammals, substantial differences are associated with processing in the Golgi complex. In the endoplasmic reticulum, both systems add core units composed of mannose, *N*-acetylglucosamine, and glucose residues, and processing of certain glucose and mannose residues follows. However, in the Golgi complex, mammalian cells continue to process mannose residues and additional residues such as galactose and sialic acid. In contrast, *S. cerevisiae* performs neither of these functions, but instead adds extensive outer chains with many branches that can be heterogeneous in size.

Genetic approaches have been employed to facilitate the production in yeast of glycoproteins more similar to their human counterparts. Ballou and colleagues (Ballou et al., 1980) provided a basis for this approach based on careful genetic dissection of the glycosylation pathway. For example, yeast strains with the *mnn9* mutation do not add further mannose units in the Golgi complex. In the case of At-III, use of strains carrying this mutation resulted in production of a homogeneous protein with only core glycosylation. Further utilization of the *mnn* family of mutations may permit engineering of yeast-produced glycoproteins with improved properties.

Expression levels can markedly affect post-translational modification of foreign proteins. Thus, in the At-III example, expression at relatively low levels resulted in protein that carried full glycosylation, including inner and outer chains. However, as expression was enhanced, glycosylation was reduced, and at high-level expression directed by YpGX262, most of the At-III was unglycosylated. This results, of course, from the important interplay that occurs between the expression level and

BRUNEL UNIVERSITY LIBRARY

the ability of yeast cells to process the molecules during their transport through the secretory pathway.

Secretion to the Extracellular Medium

Although some success has been obtained in secreting complex mammalian proteins extracellularly from *S. cerevisiae*, the greatest success has only been achieved with smaller peptides or proteins normally secreted by fungal sources. Indeed, in the At-III example, no extracellular protein was detected. Other yeasts appear to secrete mammalian proteins more effectively, posing a major challenge to the future of *S. cerevisiae* as the system of choice for secreted proteins. However, increased understanding of yeast genetics and biochemistry may be used to substantially improve secretion of foreign proteins in *S. cerevisiae*.

Transport of proteins through a eukaryotic secretory pathway is a complex process involving the participation of many cofactors. For instance, signal peptidase activity is required for removal of signal sequences from mature proteins. In the case of α-factor, additional proteases encoded by the *KEX2* and *STE13* genes are required for cleavage of the pro sequence from the mature peptide. Thus, it is possible that with overexpression of secreted proteins, these cofactors become limiting. This is supported by the observation of Haguenauer-Tsapis and Hinnen (1984) that upon overexpression of yeast acid phosphatase, precursor protein accumulates. In some cases, overexpression of limited processing enzymes has a beneficial effect. Barr et al. (1987) noted that multiple copies of *KEX2* enhanced the ability of *S. cerevisiae* to process and secrete a foreign protein, transforming growth factor, linked to the α-factor prepro sequence. Therefore, careful examination of processing defects can identify limited processing enzymes within the cell and their increased expression can, at least in some cases, improve processing and secretion.

In addition to protease activities, many other protein functions are required for efficient secretion. The activity of several accessory proteins is required for proper conformational folding within the endoplasmic reticulum—e.g., molecular chaperones as well as proteins specifically associated with disulfide bond formation. Robinson et al. (1994) reported that overexpression of protein disulfide isomerase, an enzyme that catalyzes disulfide exchange within the endoplasmic reticulum, resulted in improved secretion of both human platelet-derived growth factor B homodimer and *Schizosaccharomyces pombe* acid phosphatase from *S. cerevisiae*.

Furthermore, the ease of *S. cerevisiae* genetic manipulation offers the potential for improved secretion even in cases where specific defects have not been identified. Toward that end, several groups have applied mutagenesis strategies in conjunction with rapid screening methods to identify yeast strains with enhanced secretion capability. Smith et al. (1985) applied this approach to identify yeast strains that carried mutations leading to enhanced secretion of chymosin. Importantly, improved secretion of additional, but not all, foreign proteins was associated with one of these mutations.

TIME FRAME AND FUTURE PROSPECTS

As summarized in this commentary, tools for the expression of foreign genes in *S. cerevisiae* are very well developed, and allow the individual researcher to establish within a few months the relative capabilities of *S. cerevisiae* as a host for expression of new target proteins. However, in many cases the challenges are substantial and long-term dedicated efforts are required for ultimate success. Prospects for further engineering of *S. cerevisiae* cells for the expression of foreign proteins seem quite bright, in part because a cooperative international effort is likely to soon complete the entire sequence of the *S. cerevisiae* genome. The availability of a complete repertoire of genetic information will provide a very firm basis for better understanding of complex eukaryotic pathways, and will likely suggest new avenues for engineering this organism for enhanced expression of complex mammalian proteins.

LITERATURE CITED

Ballou, L., Cohen, R., and Ballou, C. 1980. *Saccharomyces cerevisiae* mutants that make mannoproteins with a truncated carbohydrate outer chain. *J. Biol. Chem.* 255:5986-5991.

Barr, P.J., Steimer, K.S., Sabin, E.A., Parkes, D., George-Nascimento, C., Stephans, J.C., Powers, M.A., Gyenes, A., Nan Nest, G.A., Miller, E.T., Higgins, K.W., and Luciw, P.A. 1987. Antigenicity and immunogenicity of domains of the human immunodeficiency virus (HIV) envelope protein expressed in the yeast *Saccharomyces cerevisiae*. *Vaccine* 5:90-101.

Beggs, J.D. 1978. Transformation of yeast by a replicating hybrid plasmid. *Nature* 275:104-109.

Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., and Barr, P.J. 1984. α-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cere-*

UCLA CENTRAL UNIVERSITY LIBRARY

visiae. Proc. Natl. Acad. Sci. U.S.A. 81:4642-4646.

Castelli, L.A., Mardon, C.J., Strike, P.M., Azad, A.A., and Macreadie, I.G. 1994. High-level secretion of correctly processed beta-lactamase from *Saccharomyces cerevisiae* using a high-copy-number secretion vector. *Gene* 142:113-117.

Erhart, E. and Hollenberg, C.P. 1983. The presence of a defective *LEU2* gene in 2μm DNA recombinant plasmids is responsible for curing and high copy number. *J. Bacteriol.* 156: 625-635.

Etcheverry, T. 1990. Induced expression using yeast copper metallothionein promoter. *Methods Enzymol.* 185:319-329.

Haguenauer-Tsapis, R. and Hinnen, A. 1984. A deletion that includes the signal peptidase cleavage site impairs processing, glycosylation, and secretion of cell-surface yeast acid phosphatase. *Mol. Cell. Biol.* 4:2668-2675.

Henderson, R.C.A., Cox, B.S., and Tubb, R. 1985. The transformation of brewing yeasts with a plasmid containing the gene for copper resistance. *Curr. Genet.* 9:133-138.

Hinnen, A., Meyhack, B., and Heim, J. 1989. Heterologous gene expression in yeast. *In* Yeast Genetic Engineering (P.J. Barr, A.J. Brake, and P. Valenzuela, eds.) pp. 193-213. Butterworths, London.

Hitzeman, R.A., Chen, C.Y., Dowbenko, D.J., Renz, M.E., Lui, C., Pai, R., Simpson, N.J., Kohr, W.J., Singh, A., Chisholm, V., Hamilton, R., and Chang, C.N. 1990. Use of heterologous and homologous signal sequences for secretion of heterologous proteins from yeast. *Methods Enzymol.* 185:421-440.

Innis, M.A., Holland, M.J., McCabe, P.C., Cole, G.E., Wittman, V.P., Tal, R., Watt, K.W.K., Gelfand, D.H., Holland, J.P., and Meade, J.H. 1985. Expression, glycosylation, and secretion of an *Aspergillus* glucoamylase by *Saccharomyces cerevisiae*. *Science* 228:21-26.

Johnston, M. 1987. A model fungal gene regulatory mechanism: The *GAL* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* 51:458-47.

Kingsman, S.M., Kingsman, A.J., Dobson, M.J., Mellor, J., and Roberts, N.A. 1985. Heterologous gene expression in *Saccharomyces cerevisiae*. *Biotechnol. Genet. Eng. Revs.* 3:377-416.

Kramer, R.A., DeChiara, T.M., Schaber, M.D., and Hilliker, S. 1984. Regulated expression of a human interferon gene in yeast: Control by phosphate concentration or temperature. *Proc. Natl. Acad. Sci. U.S.A.* 81:367-370.

Kukuruzinska, M.A., Bergh, M.L.E., and Jackson, B.L. 1987. Protein glycosylation in yeast. *Annu. Rev. Biochem.* 56:915-944.

Lemontt, J.F., Wei, C.M., and Dackowski, W.R. 1985. Expression of active human tissue plasminogen activator in yeast. *DNA* 4: 419-428.

Lopes, T.S., Klootwijk, J., Veenstra, A.E., van der Aar, P.C., van Heerikhuizen, H., Raue, H.A., and Planta, R.J. 1989. High-copy-number integration into the ribosomal DNA of *Saccharomyces cerevisiae*: A new vector for high-level expression. *Gene* 79:199-206.

Penttila, M.E., André, L., Lehtovaara, P., Bailey, M., Teeri, T.T., and Knowles, J.K.C. 1988. Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. *Gene* 63:103-112.

Robinson, A.S., Hines, V., and Wittrup, K.D. 1994. Protein disulfide isomerase overexpression increases secretion of foreign proteins in *Saccharomyces cerevisiae*. *Bio/Technology* 12:381-384.

Romanos, M.A., Scorer, C.A., and Clare, J.J. 1992. Foreign gene expression in yeast: A review. *Yeast* 8:423-488.

Sakai, A., Ozawa, F., Higashizaki, T., Shimizu, Y., and Hishinuma, F. 1991. Enhanced secretion of human nerve growth factor from *Saccharomyces cerevisiae* using an advanced δ-integration system. *Bio/Technology* 9:1382-1385.

Shuster, J.R. 1989. Regulated transcriptional systems for the production of proteins in yeast: Regulation by carbon source. *In* Yeast Genetic Engineering (P.J. Barr, A.J. Brake, and P. Valenzuela, eds.) pp. 83-108. Butterworths, London.

Smith, R.A., Duncan, M.J., and Moir, D.T. 1985. Heterologous protein secretion from yeast. *Science* 230:1219-1224.

Strausberg, R.L. and Strausberg, S.L. 1991. Composite yeast vectors. U. S. Patent 5,013,652.

Tuite, M.F., Dobson, M.J., Roberts, N.A., King, R.M., Burke, D.C., Kingsman, S.M., and Kingsman, A.J. 1982. Regulated high-efficiency expression of human interferon-α in *Saccharomyces cerevisiae*. *EMBO J.* 1:603-608.

Van Arsdell, J.N., Kwok, S., Schweikart, V.L., Ladner, M.B., Gelfand, D.H., and Innis, M.A. 1987. Cloning, characterization, and expression in *Saccharomyces cerevisiae* of endoglucanase I from *Trichoderma reesei*. *Bio/Technology* 5:60-64.

Wanker, E., Klingsbichel, E., and Schwab, H. 1995. Efficient secretion of *Bacillus subtilis* levanase by *Saccharomyces cerevisiae*. *Gene* 161:45-49.

Webster, T.D. and Dickson, R.C. 1983. Direct selection of *Saccharomyces cerevisiae* resistant to the antibiotic G418 following transformation with a DNA vector carrying the kanamycin resistance gene of Tn *903*. *Gene* 26:243-252.

Contributed by Robert L. Strausberg
National Institutes of Health
Bethesda, Maryland

Susan L. Strausberg
University of Maryland
Rockville, Maryland

BOWMAN-GUNTER-WILSON LIBRARY

Overview of Protein Expression in *Pichia pastoris*

During the early 1970s, the methylotrophic yeast *Pichia pastoris* was developed as a biological tool to convert methanol—a cheap and readily available carbon source that was being discarded by the oil industry as a waste product—into high-quality protein that could be used in the livestock industry. A combination of the oil crisis that began in the mid-1970s, the fact that *Pichia* as a component of livestock feed was never able to compete in cost with soybeans, and the emergence of the biotechnology industry led to new attempts to find potential uses for the high protein-production capacity of *Pichia*. By the early 1980s the focus had shifted to the use of *Pichia* as a eukaryotic microbial system to produce large quantities of heterologous proteins of interest to biotechnologists, including those in the pharmaceutical industry (Ratner, 1989; Wegner, 1990; Buckholz et al., 1991; Cregg et al., 1993). The successful and rapid introduction of *P. pastoris* into the general scientific community as a heterologous expression system is a result of several important factors, including the goal of developing *P. pastoris* into a single-cell protein (SCP) production system, its utility as a model eukaryotic organism for basic molecular and cellular biology studies (Gould et al., 1992), and its similarity—both biologically and technically—to one of the most well-characterized systems in modern biology, the baker's yeast *Saccharomyces cerevisiae* (UNIT 5.6).

The exploitation of *Pichia* as a tool to produce useful quantities of biologically relevant proteins has moved into high gear (see Table 5.7.1). *Pichia* as an expression system is now readily available to the research community and can be used to generate proteins destined for academic scrutiny in addition to proteins produced in *Pichia* for commercialization. The number of heterologous proteins successfully being expressed in *Pichia* is doubling almost monthly.

GENERAL CHARACTERISTICS OF *PICHIA PASTORIS*

As a yeast, *Pichia pastoris* is a microbial eukaryote and is as easy to manipulate as *Escherichia coli*. It has many of the advantages of eukaryotic expression (e.g., protein processing, folding, and post-translational modifications), and it is faster, easier, and cheaper to use than other eukaryotic expression systems, such as baculovirus (UNITS 5.4 & 5.4) or mammalian tissue culture. It also generally gives higher expression levels.

P. pastoris is completely amenable to the genetic, biochemical, and molecular biological techniques that have been developed over the past several decades for *S. cerevisiae* with little or no modification. In particular, methods for transformation by complementation, gene disruption, and gene replacement developed for *S. cerevisiae* work equally well for *P. pastoris* (Cregg et al., 1987, 1989; Guthrie and Fink, 1991). In addition, heterologous protein expression levels are often 10- to 100-fold higher in *Pichia*.

The genetic nomenclature adopted for *P. pastoris* mirrors that used for *S. cerevisiae*. For example, the gene from *S. cerevisiae* that encodes the enzyme histidinol dehydrogenase is called the *HIS4* gene; the homologous gene from *P. pastoris* is called the *P. pastoris HIS4* gene, and so on. There is also a very high degree of cross-functionality between *P. pastoris* and *S. cerevisiae*. For instance, many *S. cerevisiae* genes have been shown to genetically complement the comparable mutants in *P. pastoris*, and vice versa—e.g., the *P. pastoris HIS4* gene functionally complements *S. cerevisiae his4* mutants, and the *S. cerevisiae HIS4* gene functionally complements *P. pastoris his4* mutants. Other cross-complementing genes that have been identified include *LEU2*, *ARG4*, *TRP1*, and *URA3*.

P. pastoris as a Methylotrophic Yeast

P. pastoris, one of four different genera of methylotrophic yeasts (the others being *Candida*, *Hansenula*, and *Torulopsis*), is capable of metabolizing methanol as its sole carbon source. The first step in this process is the oxidation of methanol to formaldehyde by the enzyme alcohol oxidase, generating hydrogen peroxide in the process. To avoid hydrogen peroxide toxicity, methanol metabolism takes place within a specialized cell organelle, the peroxisome, which sequesters toxic intermediates from the rest of the cell. Alcohol oxidase is sequestered in the peroxisomes and functions as a homo-octomer, with each subunit containing one noncovalently bound flavin adenine dinucleotide (FAD) cofactor. Alcohol oxidase

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

Table 5.7.1 Examples of Heterologous Protein Expression in *Pichia pastoris*

Protein	Expression level (g/liter)	Where/how expressed[a]	Reference
Hepatitis B surface antigen	0.4	I/Mut^S	Cregg et al. (1987)
β-Galactosidase	20,000 (U/mg total protein)	I/Mut$^+$	Tschopp et al. (1987a)
Invertase	2.3	S/Mut$^+$	Tschopp et al. (1987b)
Tumor necrosis factor (TNF)	10.0	I/Mut^S	Sreekrishna et al. (1989)
Bovine lysozyme *c*2	0.55	S/Mut$^+$	Digan et al. (1989)
Streptokinase (active)	0.08	I/n.r.	Hagenson et al. (1989)
Tetanus toxin fragment C	12.0	I/Mut$^+$ and Mut^S	Clare et al. (1991a)
Pertussis antigen P69	3.0	I/Mut^S	Romanos et al. (1991)
Mouse epidermal growth factor (EGF)	0.45	S/Mut^S	Clare et al. (1991b)
Aprotinin	0.8	S/Mut$^+$/Mut^S	Vedvick et al. (1991)
Kunitz protease inhibitor	1.0	S/Mut$^+$	Wagner et al. (1992)
Human serum albumin	4.0	S/Mut$^+$ and Mut^S	Barr et al. (1992)
HIV-1 gp120	1.25	I and S/Mut$^+$	Scorer et al. (1993)
Carboxypeptidase B	0.8	S/Mut$^+$ and Mut^S	Despreaux and Manning (1993)
TAP (tick anticoagulant protein)	1.7	S/Mut$^+$	Larouche et al. (1994)
Porcine leukocyte lipoxygenase (72 kD)	n.r.	I/Mut^S	Reddy et al. (1994)
Human proteinase cathepsin E (homodimer of 42 kD protein)	n.r.	S/Mut^S	Yamada et al. (1994)
Bm86 (tick gut glycoprotein)	1.5	S/Mut^S	Rodriguez et al. (1994)
Alpha amylase	2.5	S/Mut^S	Paifer et al. (1994)
Kunitz protease inhibitor (APLP-2)	1.0	S/Mut$^+$	Van Nostrand et al. (1994)
Thrombomodulin	0.01-0.03	S/Mut^S	White et al. (1994)
Rabbit scFv (30 kD)	>0.1	S/Mut^S	Ridder et al. (1995)
CP15 (*Cryptosporidium parvum*)	n.r.	n.r.	Jenkins and Fayer (1995)
Rat Pit-1	n.r.	n.r.	Howard and Maurer (1995)
Human N-terminal transferrin	0.27	S/Mut^S	Steinlein et al. (1995)
Amyloid precursor protein	0.001-0.004	S/Mut^S	Ohsawa et al. (1995)
Human CD38	~0.16	S/Mut^S	Fryxell et al. (1995)
Hamster prion protein PrPc	<0.001	I/Mut^S	Weiss et al. (1995)
Mouse OB gene (leptin)	0.01	S/Mut$^+$	Halaas et al. (1995)
Human alanyl-tRNA synthetase	3% (total protein)	I/Mut^S	Shiba et al. (1995)
Recombinant enterokinase (26 kD)	0.006	S/Mut$^+$	Vozza et al., in press

[a]Abbreviations: I, intracellular; Mut$^+$, methanol utilization wild type (genotype *AOX1 AOX2*); Mut^S, methanol utilization slow (genotype *aox1 AOX2*); n.r., not reported or not available; S, secreted.

PCU IAS / CENTRAL UNIVERSITY LIBRARY

has a poor affinity for O_2 and *P. pastoris* compensates for this by generating large amounts of the enzyme.

There are two genes in *P. pastoris* that code for alcohol oxidase—*AOX1* and *AOX2*—but the *AOX1* gene is responsible for the vast majority of alcohol oxidase activity in the cell. Expression of the *AOX1* gene is tightly regulated and is induced by methanol to very high levels, typically ≥30% of the total soluble protein in cells grown with methanol as the carbon source. The *AOX1* gene has been isolated and a plasmid-borne version of its promoter is used to drive expression of the gene of interest encoding a desired heterologous protein (Ellis et al., 1985; Tschopp et al., 1987a; Koutz et al., 1989). Expression of the *AOX1* gene is controlled at the level of transcription. In methanol-grown cells ~5% of the polyA$^+$ RNA is from the *AOX1* gene. Regulation of the *AOX1* gene is similar to regulation of the *GAL1* gene (and others) of *S. cerevisiae* in that control involves a two-step process: a repression/derepression mechanism plus an induction mechanism. However, unlike the situation in *S. cerevisiae*, derepression alone of the *AOX1* gene (i.e., absence of a repressing carbon source such as glucose) is not sufficient to generate even minute levels of expression from the *AOX1* gene. The inducer, methanol, is also necessary to obtain even minimally detectable levels of *AOX1* expression (Ellis et al., 1985; Tschopp et al., 1987a; Koutz et al., 1989).

Loss of the *AOX1* gene (see Transformation by Integration), and thus a loss in most of the cell's alcohol oxidase activity, results in a strain that is phenotypically Mut^S (methanol utilization slow). This phenotype is caused by the loss of the alcohol oxidase activity encoded by the *AOX1* gene, which results in the cell's inability to metabolize methanol. Thus, a poor growth phenotype on methanol media is observed. (Mut$^+$ strains are wild type and capable of metabolizing methanol as their sole carbon source.)

HETEROLOGOUS PROTEIN EXPRESSION

P. pastoris has been used successfully to express a wide range of heterologous proteins as either intracellular or secreted proteins (see Table 5.7.1). Secretion requires the presence of a signal sequence on the expressed protein to target it to the secretory pathway. Although several different secretion signal sequences have been used successfully, including the native secretion signal present on some heterologous proteins, success has been variable. The secretion signal sequence from the *S. cerevisiae MATα* factor prepro peptide (*MF*-α1) has been used with the most success (Cregg et al., 1993; Scorer et al., 1993).

The major advantage of expressing heterologous proteins as secreted proteins is that *P. pastoris* secretes very low levels of native proteins. That, combined with the very low amount of protein in the minimal *Pichia* growth medium, means that the secreted heterologous protein comprises the vast majority of the total protein in the medium, and this serves as a first step in purification of the expressed protein (Barr et al., 1992).

Strains for Expression

The *P. pastoris* host strains have defects in either the histidinol dehydrogenase activity encoded by the *HIS4* gene (strains GS115, KM71, SMD1163, SMD1165, and SMD1168), or the arginosuccinate lyase activity encoded by the *ARG4* gene (strain GS190), or both (strain PPF1; see Table 5.7.2). All of these strains will grow on complex media such as YPD (yeast

Table 5.7.2 Some Strains of *Pichia pastoris* Used for Heterologous Protein Expression

Strain	Genotype	Feature	Vector selectable marker required
GS115	*his4*	Wild-type strain	*HIS4*
GS190	*arg4*	Wild-type strain	*ARG4*
KM71	*his4 arg4 aox1::ARG4*	Mut^S	*HIS4*
PPF1	*his4 arg4*	Two selectable auxotrophic markers	*HIS4* or *ARG4*
SMD1163	*his4 pep4 prB1*	Double protease mutant	*HIS4*
SMD1165	*his4 prB1*	Protease mutant	*HIS4*
SMD1168	*his4 pep4*	Protease mutant	*HIS4*

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 5.7.3 Media for *Pichia pastoris*

Media[a]	YNB[b]	$(NH_4)_2SO_4$	Peptone	Yeast extract	Carbon source[c]	Phosphate buffer[d]	Other	Use
YPD	—	—	2%	1%	2% glu	—	—	Rich medium
RDB[e]	0.34%	1%	—	—	2% glu	—	1 M sorbitol	Transformation
MD[f]	0.34%	1%	—	—	2% glu	—	—	Mut^+/Mut^S
MM[f]	0.34%	1%	—	—	0.5% meth	—	—	Mut^+/Mut^S
MGY	0.34%	1%	—	—	1% gly	—	—	Expression
MM	0.34%	1%	—	—	0.5% meth	—	—	Expression
BMGY[g]	0.34%	1%	2%	1%	1% gly	100 mM	—	Alternate expression
BMMY	0.34%	1%	2%	1%	0.5% meth	100 mM	—	Alternate expression

[a]All of these media can be made either as liquid or, by adding agarose to 2% (w/v), as plates.

[b]Yeast nitrogen base without amino acids and without $(NH_4)_2SO_4$.

[c]Abbreviations: glu, glucose (same as dextrose); gly, glycerol; meth, methanol.

[d]Stock: 1 M potassium phosphate buffer, adjusted to desired pH between 5 and 8.

[e]No amino acid supplements; these media are normally used with transformed strains that are His^+ (or Arg^+) prototrophs.

[f]MD and MM media can be used to identify Mut^S clones. Growth of both Mut^+ and Mut^S clones will be wild type on MD medium. Growth of Mut^+ will be wild type on MM medium, but growth of Mut^S on MM will be severely retarded to completely absent after 2 days because of the inability of Mut^S clones to efficiently metabolize methanol as a carbon source.

[g]The addition of yeast extract and peptone to expression media may improve the results of heterologous protein expression in some cases. The mode of action of this improvement is not known, but it may be that yeast extract and peptone serve as competitive inhibitors of proteases and/or act as weak medium buffering agents.

extract, peptone, dextrose) and on minimal media, such as MD (minimal dextrose), MM (minimal methanol), and MGY (minimal glycerol), that have been supplemented with histidine and/or arginine (Table 5.7.3). Until transformed with the appropriate vectors, these strains will not grow on minimal medium lacking histidine (those strains with genotype including *his4*) or arginine (those strains with genotype including *arg4*). Selection of His^+ or Arg^+ prototrophs among *his4/arg4* mutant strains transformed with a plasmid carrying the wild-type gene is used as the basis for selecting recombinant strains containing the expression plasmid.

Most expression experiments begin with wild-type strain GS115 (*his4*) in which vectors carrying the *HIS4* gene can be selected, but there are several alternative strains, mostly derivatives of GS115, that can be employed to achieve success empirically or when a specific need can be predicted. Both Mut^+ and Mut^S as well as single- and multicopy vector recombinant versions of GS115 can be generated.

KM71 (*his4 arg4 aox1::ARG4*) is a Mut^S strain because the chromosomal *AOX1* gene has been replaced with the *ARG4* gene. It can be used when only Mut^S recombinant strains are desired (see Table 5.7.2).

SMD1163 (*his4 pep4 prB1*), SMD1165 (*his4 prB1*), and SMD1168 (*his4 pep4*) are a series of protease-deficient strains that may provide a more suitable environment for certain proteins that are being expressed as heterologous proteins. The *PEP4* gene codes for proteinase A, a vacuolar aspartyl protease required for activation of other vacuolar proteases such as carboxypeptidase Y and proteinase B. Prior to processing and activation by proteinase A, proteinase B has about half of its protease activity. The *PRB1* gene codes for proteinase B itself. Therefore, *pep4* mutants display a substantial decrease in or elimination of proteinase A and carboxypeptidase Y activity, and partial reduction in proteinase B activity. *prB1* mutants show a substantial decrease in or elimination of proteinase B activity only, and *pep4 prB1* double mutants show a substantial reduction in or elimination of all three of these protease activities (see Table 5.7.4).

Expression Plasmids

Plasmid vectors designed for heterologous protein expression in *P. pastoris* all have general features in common. These are illustrated in the generic plasmid depicted in Figure 5.7.1. (1) The $5'P_{AOX1}$ is the promoter from the *Pichia* alcohol oxidase (*AOX1*) gene; it is used to drive

CSU MORGAN LIBRARY

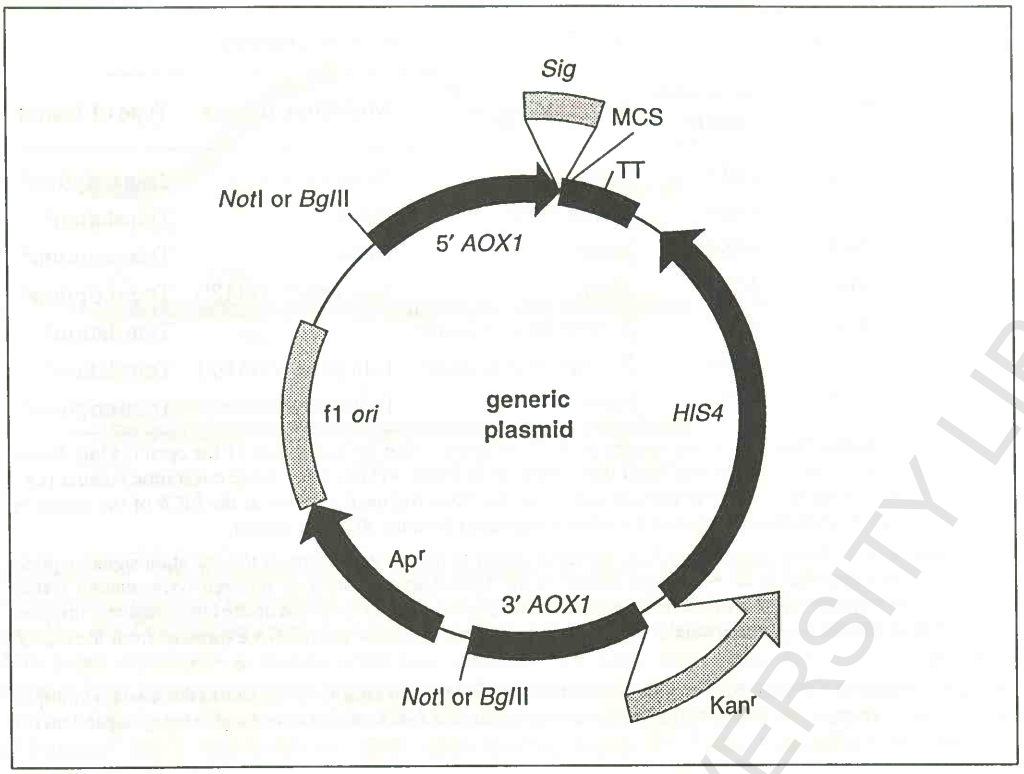


Figure 5.7.1 Generic plasmid for heterologous protein expression in *Pichia pastoris*. Elements present in all expression plasmids are shown in dark boxes, optional elements present in some plasmids are shown in gray boxes. Abbreviation: TT, transcriptional termination sequence.

methanol-inducible expression of the gene of interest. (2) The MCS is a multiple cloning site with unique restriction endonuclease sites; it is used to insert the gene of interest into the plasmid. (3) Transcription termination sequences (TT) are derived from the native *Pichia AOX1* gene; they are used to promote efficient mRNA processing and polyadenylation. (4) *HIS4*, the wild-type gene for histidinol dehydrogenase, is a selectable marker; it is used to positively select for recombinant *Pichia* strains that have acquired the vector, which complements the auxotrophic *his4* mutation. (5) The 3′*AOX1* sequence is derived from a region of the native gene that lies 3′ to the transcription termination sequences; it is required for integration of vector sequence by gene replacement or gene insertion 3′ to the chromosomal *AOX1* gene. (6) The ColE1 origin of replication and ampicillin resistance gene (Ap^r) are required for replication and maintenance in bacteria. (7) Unique restriction endonuclease sites are included that allow for the generation of linear plasmids that can integrate at the *AOX1* locus by gene replacement (*Not*I and *Bgl*II). Other unique sites at 5′P_{AOX1} (*Sac*I and *Bst*XI) and *HIS4* (*Sal*I and *Stu*I) are included to allow generation of a linear plasmid that can integrate into the genome at these loci by gene insertion.

There are several additional features that are included in some expression vectors; these serve as tools for specialized functions (see Fig. 5.7.1 and Table 5.7.4). Sig is a DNA sequence juxtaposed between 5′P_{AOX1} and the MCS region; it encodes a protein secretion signal that is expressed as an N-terminal fusion to the protein of interest. This sequence directs the passage of the protein through the secretory pathway and targets proteins that carry it out of the cell. Sig sequences that have been used for expression in *Pichia* include the *S. cerevisiae* mating type α prepro sequence (MF-1α prepro; Scorer et al., 1993, 1994), the *Pichia* acid phosphatase signal sequence (*PHO*), the *S. cerevisiae* invertase signal sequence (Tschopp et al., 1987b), and several hybrid sequences (Larouche et al., 1994). Alternatively, the gene of interest may contain a sequence that encodes its own native secretion signal; that signal sequence may be functional in *Pichia* (Digan et al., 1989; Barr et al., 1992; Despreaux et al., 1993). Kan^r is a bacterial kanamycin resistance gene; it may be included to confer resistance to the drug G418. Recombinant *Pichia* strains that contain multiple copies of Kan^r are resistant to

DUKE UNIVERSITY LIBRARY

Table 5.7.4 Vectors for Heterologous Protein Expression in *Pichia pastoris*

Vector	Size (bp)	Selectable marker	Secretion signal	Multicopy feature	Type of fusion
pHIL-D2	8157	*HIS4*	None	None	Transcription[a]
pHIL-S1	8207	*HIS4*	*Pichia PHO1*	None	Translation[b]
pPIC3	7768	*HIS4*	None	None	Transcription[a]
pPIC3K	9009	*HIS4*	None	Kan gene[a] ($G418^r$)	Transcription[a]
pPIC9	8034	*HIS4*	*S. cerevisiae* α factor		Translation[b]
pPIC9K	9272	*HIS4*	*S. cerevisiae* α factor	Kan gene[c] ($G418^r$)	Translation[b]
pAO815	7837	*HIS4*	None	In vitro multimers	Transcription[a]

[a]*Transcription fusion:* Vector does not supply an ATG initiation codon for translation of the open reading frame contained within the expressed insert; insert must bring an in-frame ATG in a favorable eukaryotic context (i.e., Kozak consensus sequence); gene of interest carried on the DNA fragment inserted at the MCS of the vector is fused within the 5′ untranslated region of the mRNA expressed from the *AOX1* promoter.

[b]*Translation fusion:* Vector supplies an ATG initiation codon to initiate translation of the secretion signal peptide destined to be expressed as an N-terminal fusion to the heterologous protein of interest; open reading frame contained within insert may or may not have an in-frame ATG; gene of interest carried on the DNA fragment inserted at the MCS of the vector is fused within the open reading frame region of the mRNA expressed from the *AOX1* promoter.

[c]*Kan gene:* Bacterial Kan gene confers resistance to the drug G418; recombinant *Pichia* strains that contain multiple copies of the Kan gene can be resistant to higher concentrations of G418, with this resistance being dependent on Kan gene copy number.

higher concentrations of G418 in a manner that correlates with the Kan^r copy number. The f1 ori sequence is a bacterial single-strand origin of replication; this sequence can be included to generate single-stranded DNA in *Escherichia coli* to facilitate plasmid mutagenesis.

Transformation by Integration

As in *S. cerevisiae*, linear DNA can generate stable transformants of *P. pastoris* via homologous recombination between the transforming DNA and regions of homology within the genome (Cregg et al., 1985, 1989). Such integrants show extreme stability in the absence of selective pressure, even when present in multiple copies.

The most commonly used expression vectors carry the *HIS4* gene for selection and are designed to be linearized with a restriction enzyme such that His^+ recombinants can be generated by vector integration at the *his4* or *AOX1* locus (see Fig. 5.7.2). Integration events at the *AOX1* locus can occur as either gene insertions or gene replacements.

A gene replacement (omega insertion) event arises from a double cross-over between the *AOX1* promoter and 3′*AOX1* regions of the vector and genome; this results in the complete removal of the *AOX1* coding region (i.e., gene replacement; Fig. 5.7.2). The resulting phenotype is His^+ Mut^S. His^+ transformants can be readily and easily screened for their Mut phenotype, with Mut^S serving as a phenotypic indicator of integration via gene replacement at the *AOX1* locus.

Gene insertion events at the *AOX1* locus arise from a single cross-over event between the *AOX1* locus in the chromosome and any of the three *AOX1* regions on the vector (*AOX1* promoter, *AOX1* TT, or 3′ *AOX1*); these events result in the insertion of one or more copies of the vector upstream or downstream of an otherwise intact and functional native *AOX1* gene (see Fig. 5.7.3). The phenotype of such a transformant is His^+ Mut^+ (Mut^+ is wild-type for methanol utilization); the His^+ transformant is still able to utilize methanol because the *AOX1* gene remains intact and functional (see Fig. 5.7.3). Gene insertion events at the *his4* locus arise from a single cross-over event between the *his4* locus in the chromosome and the *HIS4* gene on the vector, and result in the insertion of one or more copies of the vector at the *his4* locus (see Fig. 5.7.4). Because the genomic *AOX1* locus is not involved in this recombination event, the phenotype of such a His^+ transformant remains Mut^+.

Multiple gene insertion events at a single locus in a cell do occur spontaneously, with a low but detectable frequency—between 1% and 10% of all selected His^+ transformants (see Fig. 5.7.5). Multicopy events can occur as gene

SCUI IAST CENTRAL UNIVERSITY LIBRARY

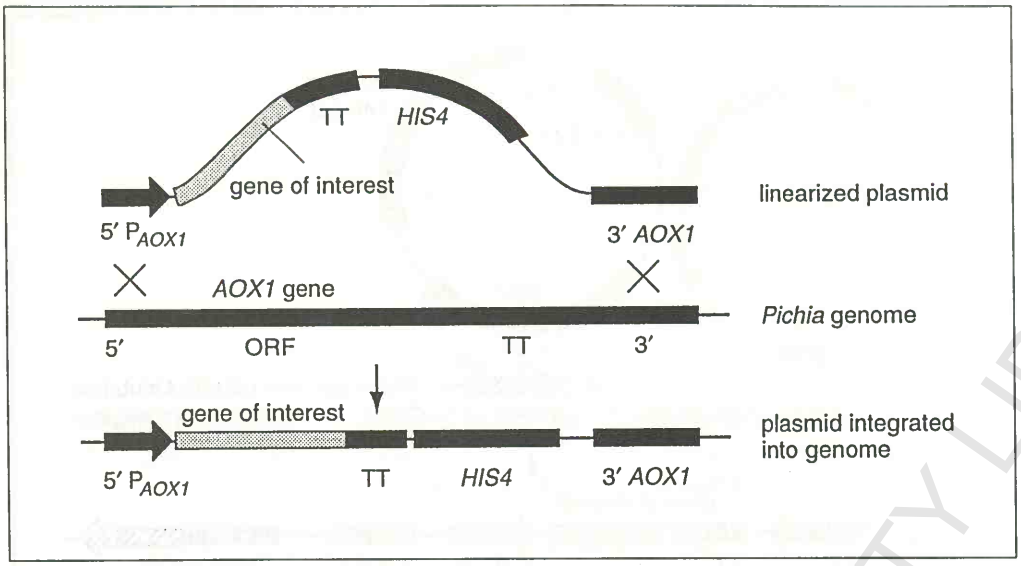


Figure 5.7.2 Gene replacement. A gene replacement event at the *AOX1* locus by a plasmid fragment carrying an "expression cassette" consisting of an *AOX1* promoter–driven gene of interest and the *HIS4* selectable marker allows for selection of His^+ prototrophs using a *his4 Pichia* strain. The event represents a double cross-over between two linear DNA molecules. Double-stranded ends of DNA molecules act to target and stimulate homologous recombination. The result is a loss of the *AOX1* locus (the strain becomes Mut^S) and the gain of P_{AOX1}-gene of interest-*HIS4* (gene of interest$^+$ His^+). Abbreviations: ORF, open reading frame; TT, transcriptional termination sequence.

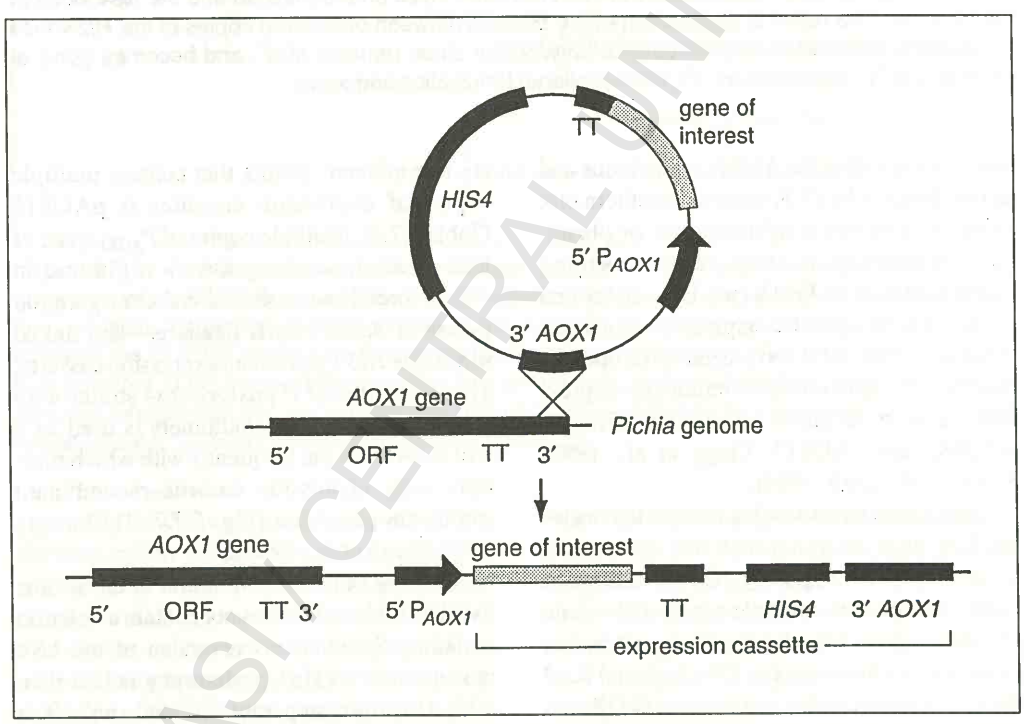


Figure 5.7.3 Gene insertion. A gene insertion event 3′ of *AOX1* in the genome by the plasmid carrying an expression cassette allows for selection of His^+ prototrophs using a *his4 Pichia* strain. The event represents a single cross-over event between circular and linear DNA molecules. Recombination takes place within regions of homology between the plasmid and the genome. The result is an insertion of the plasmid 3′ to the intact *AOX1* locus (the strain remains Mut^+) and the gain of P_{AOX1}-gene of interest-*HIS4* (gene of interest$^+$ His^+). This event also could have taken place between the 5′ *AOX1* regions on the plasmid and genome with the resulting insertion positioned 5′ to an intact *AOX1* locus. Abbreviations: ORF, open reading frame; TT, transcriptional termination sequence.

UNSW CENTRAL UNIVERSITY LIBRARY

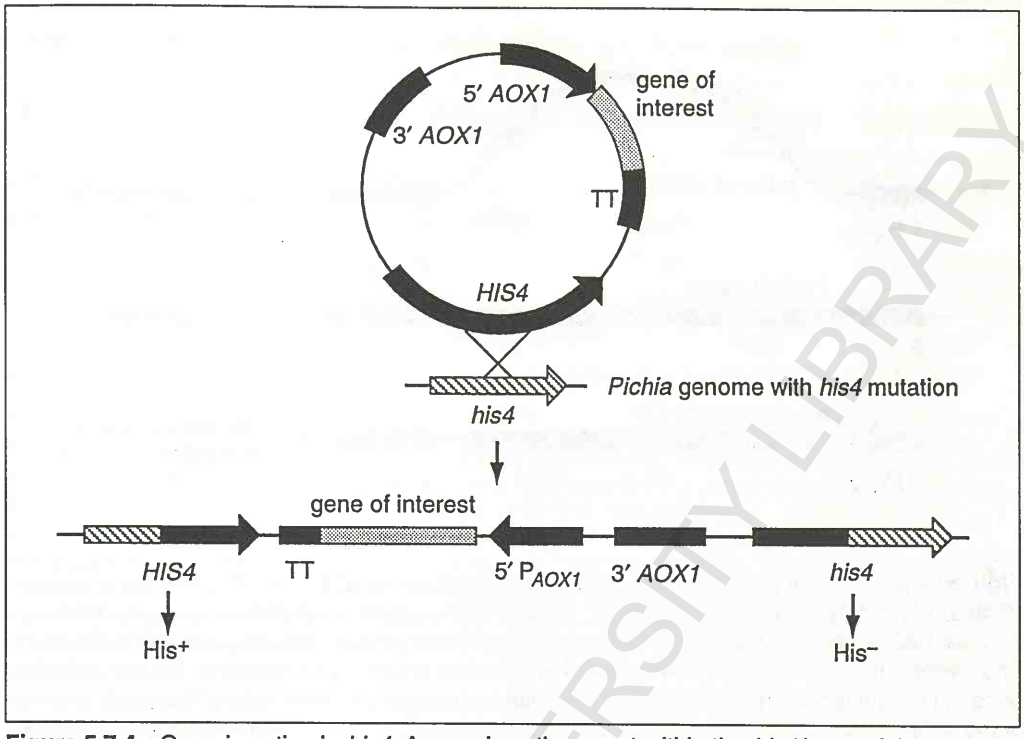


Figure 5.7.4 Gene insertion in *his4*. A gene insertion event within the *his4* locus of the genome by the plasmid carrying an expression cassette allows for selection of His$^+$ prototrophs using a *his4 Pichia* strain. The event consists of a single cross-over between circular and linear DNA molecules, in which recombination takes place between the *HIS4* locus on the plasmid and the *his4* locus in the genome. The result is an insertion of the plasmid between duplicated copies of the *HIS4/his4* genes, one still mutant and the other wild-type (the strain remains Mut$^+$, and becomes gene of interest$^+$ His$^+$). Abbreviations: TT, transcriptional termination sequence.

insertions at either the *AOX1* or *his4* locus and can be detected by PCR analysis, Southern blot analysis, differential hybridization, or phenotypic identification based on the resulting hyperresistance to G418 (see Figs. 5.7.5 and 5.7.6). Three specific expression plasmids listed in Table 5.7.4 have been developed to facilitate the generation of multicopy expression cassette recombinant strains, pPIC3K, pPIC9K, and pAO815 (Cregg et al., 1993; Scorer et al., 1993, 1994).

Expression plasmids that contain the bacterial Kan^r gene act as a marker that can be used to identify phenotypically His$^+$ recombinant strains which carry multiple copies of the gene of interest (see Fig. 5.7.6). Each expression cassette introduces a copy of the bacterial Kan^r gene. This gene confers resistance to G418, and multiple copies of the gene make recombinant strains hyperresistant to the drug G418. Hyperresistance to G418 allows for phenotypic identification of strains carrying multiple copies of the Kan^r gene and, by linkage, multiple copies of the expression cassette (P_{AOX1}-gene of interest and *HIS4*; Scorer et al., 1993, 1994).

Another plasmid system designed to generate recombinant strains that contain multiple copies of expression cassettes is pAO815 (Table 5.7.4). Multiple copies of P_{AOX1}-gene of interest are ligated together in a plasmid in vitro—forced into a directional configuration by use of *Bam*HI-*Bgl*II ligations—and linked to a single *HIS4* gene in an expression cassette. Transformation of *P. pastoris his4* strains with these in vitro–formed multimers is used as a tool to increase the frequency with which multiple copy expression cassette–recombinant strains are generated (Fig. 5.7.7; Thill et al., 1990; Cregg et al., 1993).

The *his4* mutant allele found in the strains listed in Table 5.7.2 does not contain a deletion mutation. Spontaneous reversion of the *his4* mutant allele to His$^+$ prototrophy is less than 10^{-8}. However, apparent "hit-and-run" gene conversion events of *his4* strains transformed with a *HIS4* plasmid—that is, nonintegrative recombination between the *his4* gene on the chromosome and the *HIS4* gene on the plasmid that results in conversion of the chromosomal *his4* gene to His$^+$—occurs with a frequency of 10% to 50% of all His$^+$ transformants. The His$^+$ gene conversion frequency apparently can vary

BCU IAS CENTRAL UNIVERSITY LIBRARY

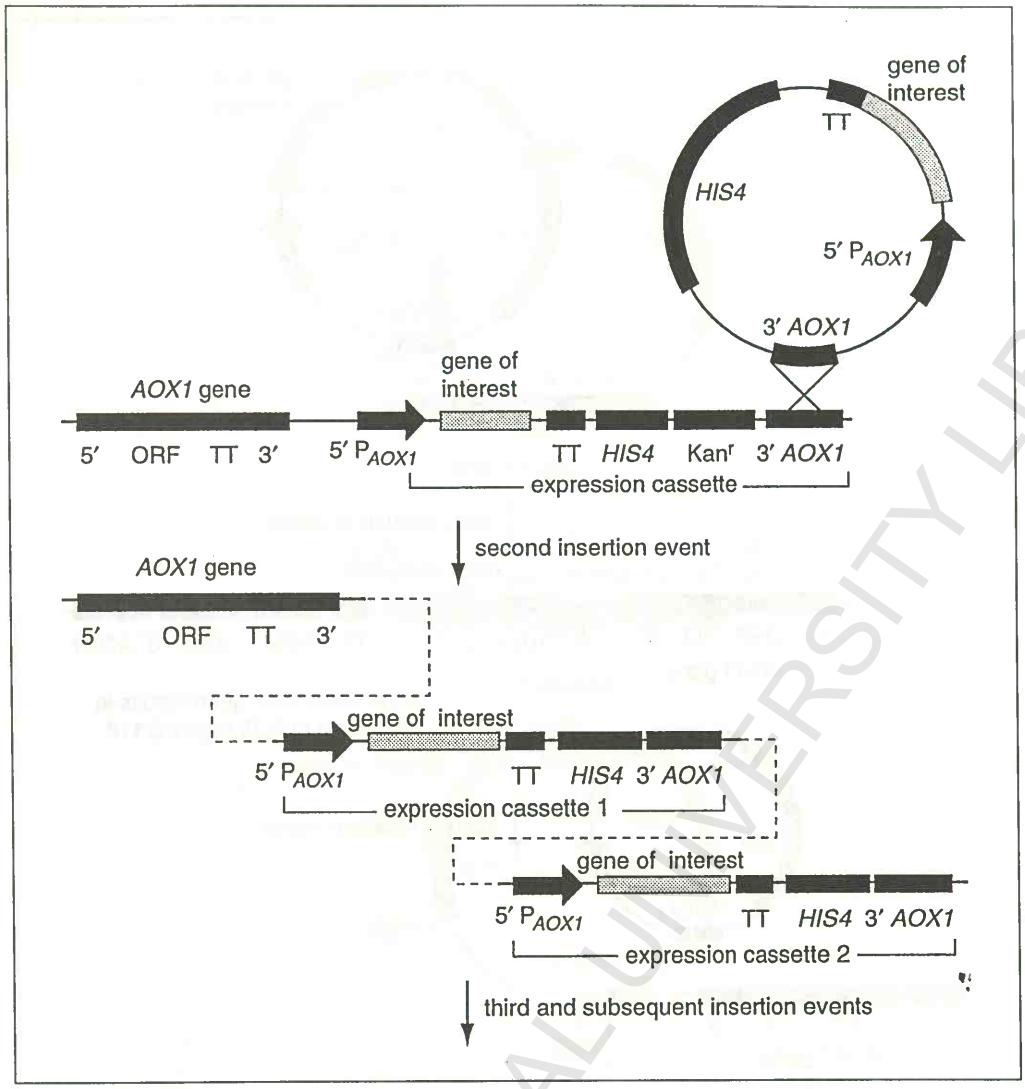


Figure 5.7.5 Multiple insertion events. Multiple insertions of an expression plasmid 3′ to an otherwise intact *AOX1* locus can result from multimer formation prior to a single insertion event or from multiple sequential insertion events at the same locus. These events can also take place between the 5′ *AOX1* regions on the plasmid and genome, with the resulting insertions positioned 5′ to an intact *AOX1* locus. Multiple events (two or more plasmids integrated) occur spontaneously, although infrequently, accounting for 1% to 10% of all His^+ transformants. Abbreviations: ORF, open reading frame; TT, transcriptional termination sequence.

depending on the transformation method, with electroporation yielding a higher frequency of His^+ gene conversion than spheroplast transformation (D.R.H., unpub. observ.). These His^+ convertants that do not contain a copy of the gene to be expressed can be identified by PCR or Southern blot analysis and eliminated from subsequent protein-expression testing.

Figure 5.7.8 outlines the steps required to generate and analyze recombinant strains of *P. pastoris*.

Post-Translational Modifications

P. pastoris utilizes most of the post-translational modification pathways typically associated with higher eukaryotes. In comparison to *S. cerevisiae*, *Pichia* may have an advantage in the glycosylation of secreted proteins because it may have less of a tendency to hyperglycosylate. Although both *S. cerevisiae* and *P. pastoris* have a majority of N-linked glycosylation of the high-mannose type, the length of the oligosaccharide chains added post-translationally to proteins is much shorter in *Pichia* than in *S. cerevisiae* (an average 8 to 14 mannose residues per side chain in *P. pastoris* compared to 50 to 150 in *S. cerevisiae*; Tschopp et al., 1987b; Grinna and Tschopp, 1989). There is also a difference between *S. cerevisiae* and *P. pastoris* in the structure of the oligosaccharides added.

B. C. U. „M. EMINESCU" IAȘI
CHITIT

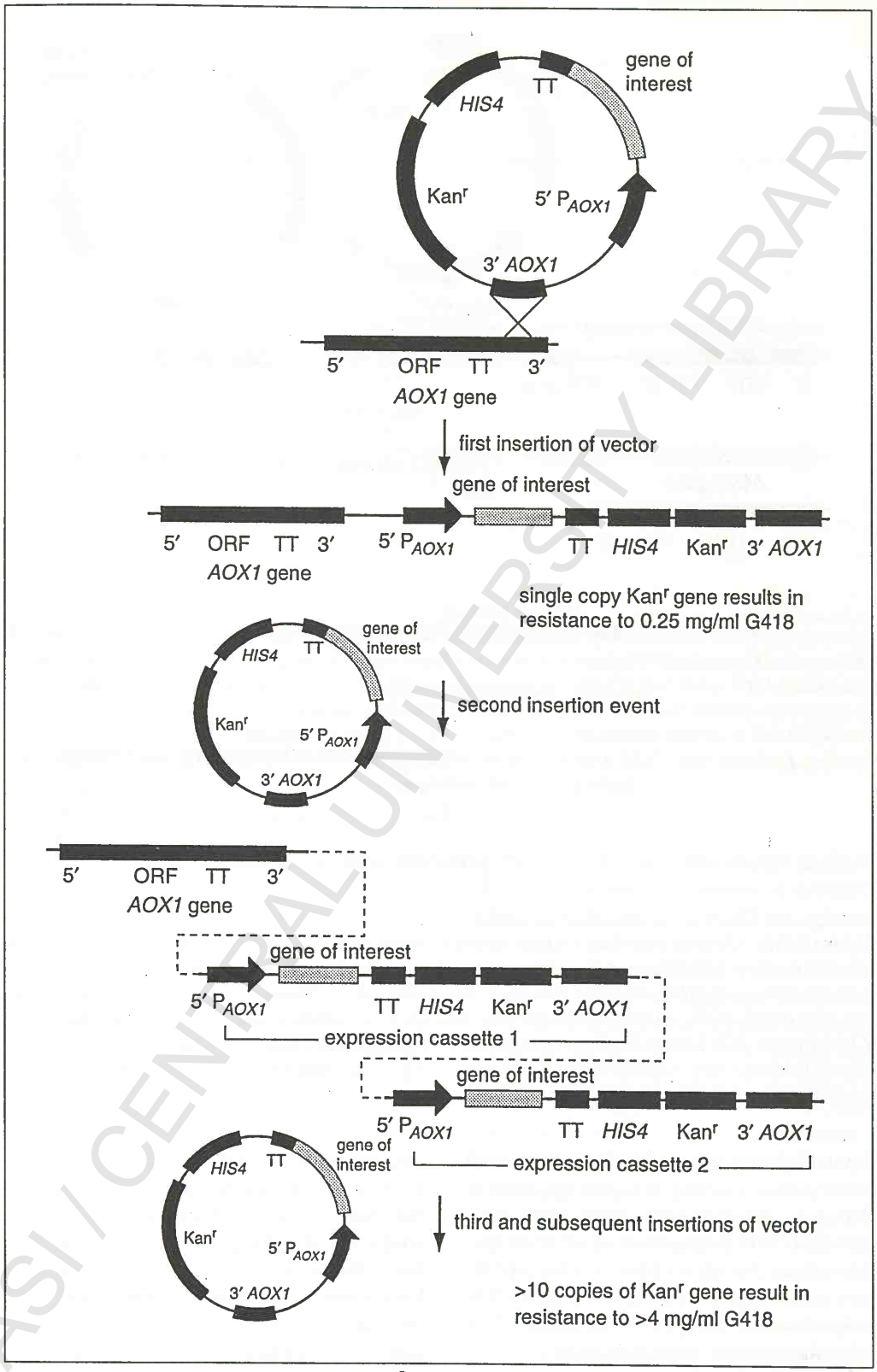


Figure 5.7.6 Multiple insertions with a Kan^r-containing plasmid. Multiple insertions of an expression plasmid 3′ to an otherwise intact *AOX1* locus can result from multimer formation prior to a single insertion event or from multiple sequential insertion events at the same locus. These events can also occur between the 5′ *AOX1* regions on the plasmid and genome, with the resulting insertions 5′ to an intact *AOX1* locus. Each expression cassette introduces a copy of the bacterial Kan^r gene, which confers resistance to G418; multiple copies make recombinant strains hyperresistant to G418, which phenotypically identifies strains carrying multiple copies of the expression cassette. Abbreviations: ORF, open reading frame; TT, transcriptional termination sequence.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

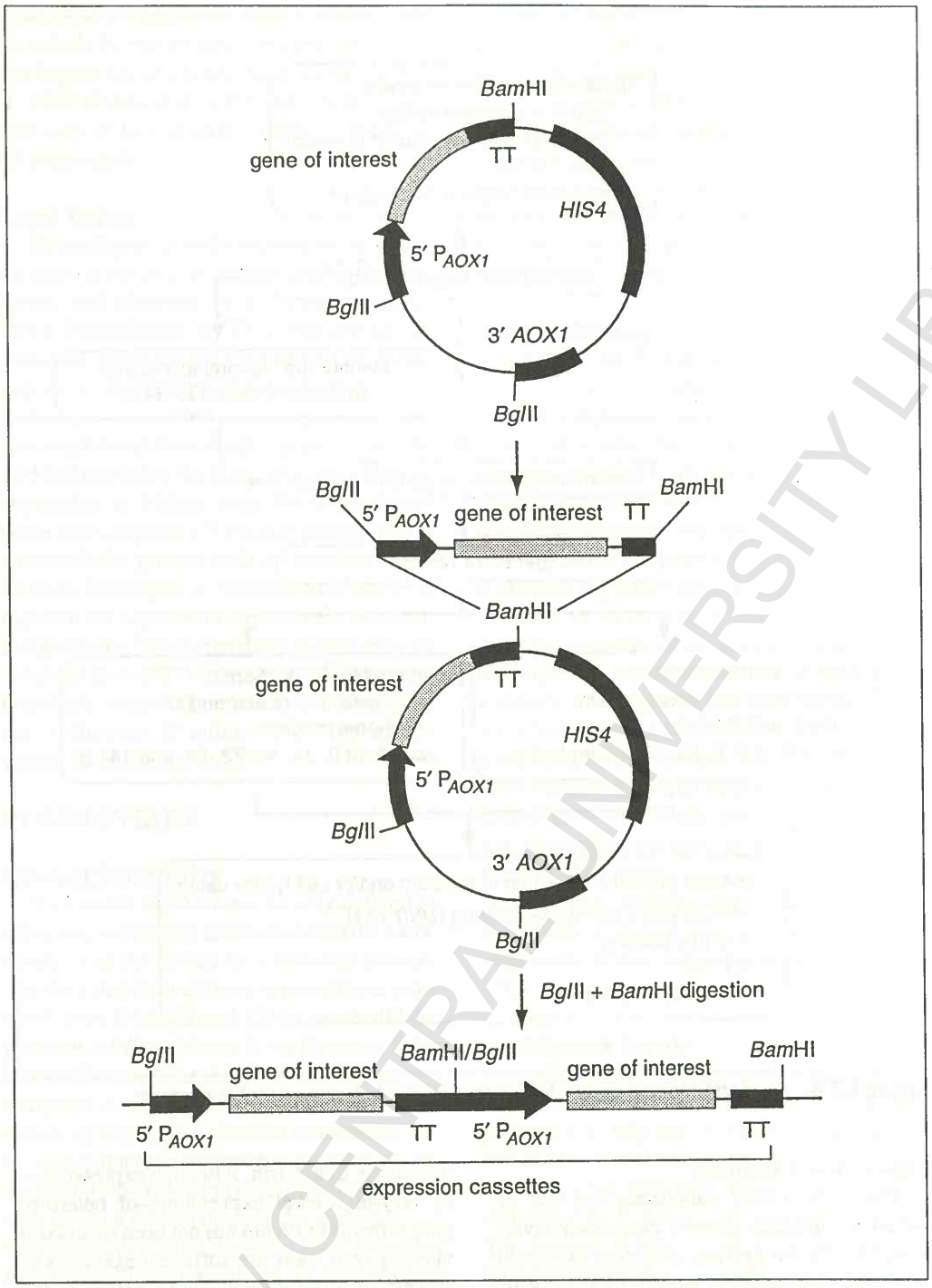


Figure 5.7.7 In vitro multimer formation. By this scheme, multiple copies of P_{AOX1}-gene of interest are ligated together in a plasmid in vitro, and linked to a single *HIS4* gene in an expression cassette. Transformation of *Pichia* with these in vitro–formed multimers is used as a tool to increase the frequency with which multiple-copy expression cassette recombinant strains are generated. Abbreviations: OEF, open reading frame; TT, transcriptional termination sequence.

S. cerevisiae core oligosaccharides have terminal α1-3 glycan linkages, whereas *P. pastoris* does not. Furthermore, it is believed that the α1-3 glycan linkages of *S. cerevisiae* are primarily responsible for the hyperantigenic nature of glycosylated proteins from *S. cerevisiae* that makes them particularly unsuitable for therapeutic use. Although not yet proven, this is predicted to be less of a problem for glycoproteins generated in *P. pastoris*, because their glycoprotein structure may resemble that of higher eukaryotes (Cregg et al., 1993).

PSU IACS CENTRAL UNIVERSITY LIBRARY

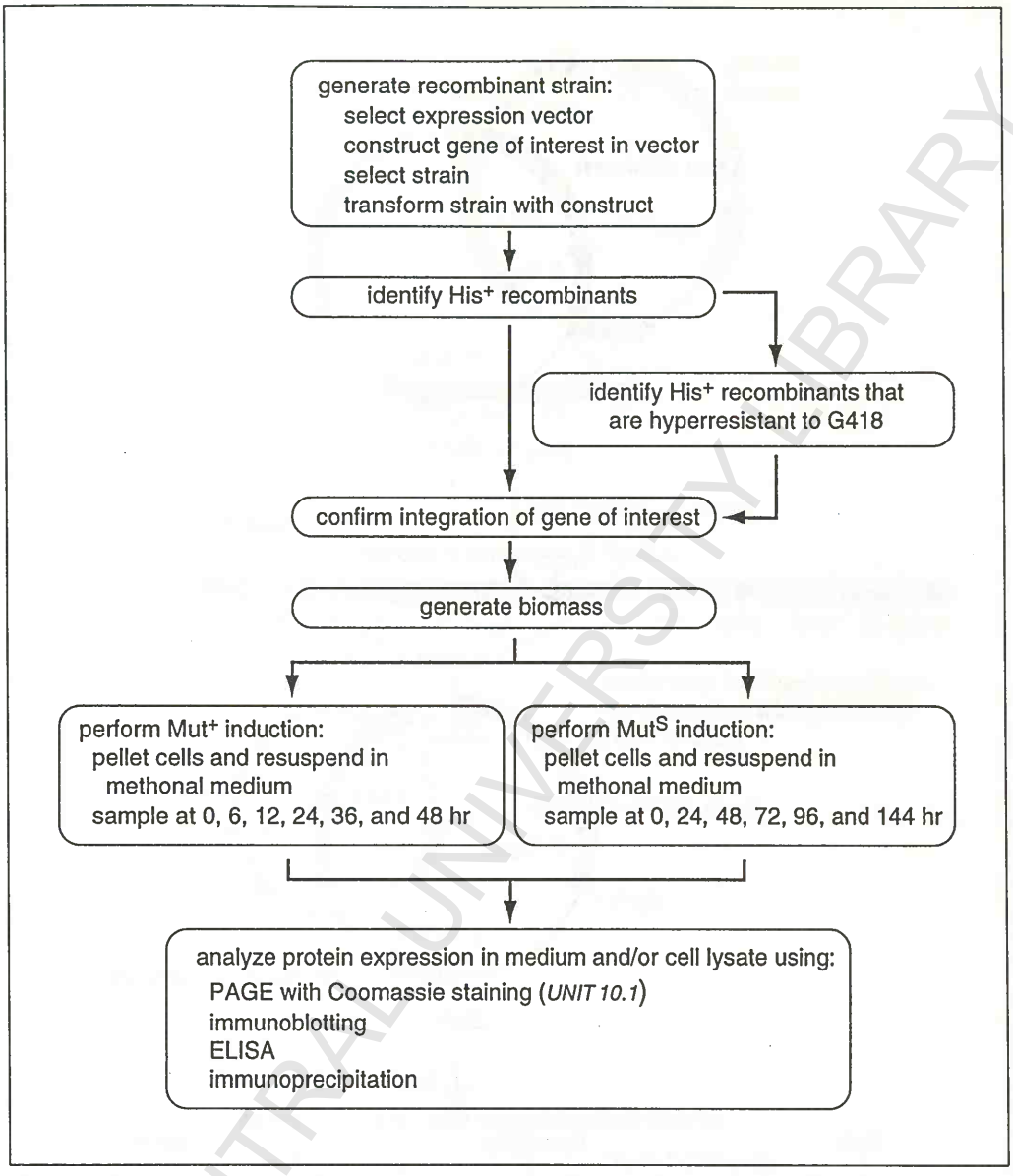

generate recombinant strain:
select expression vector
construct gene of interest in vector
select strain
transform strain with construct
identify His+ recombinants
identify His+ recombinants that are hyperresistant to G418
confirm integration of gene of interest
generate biomass
perform Mut+ induction:
pellet cells and resuspend in methonal medium
sample at 0, 6, 12, 24, 36, and 48 hr
perform MutS induction:
pellet cells and resuspend in methonal medium
sample at 0, 24, 48, 72, 96, and 144 hr
analyze protein expression in medium and/or cell lysate using:
PAGE with Coomassie staining (UNIT 10.1)
immunoblotting
ELISA
immunoprecipitation

Figure 5.7.8 Flow chart for the generation of recombinant strains of *Pichia pastoris*.

Expression Examples

There are several parameters that can be varied to optimize protein expression levels using the *Pichia* system, including the strain (e.g., protease-deficient strains, Mut$^+$ or Mut^S recombinant strains), intracellular or secreted expression, and flask or fermenter format. Table 5.7.1 lists published examples of heterologous protein expression in *Pichia*. Unpublished examples of proteins expressed in *Pichia* include MHC class II, shark cytochrome p450, muscle kinase, human serum proteins, maize heat shock protein, α-glycosidase, diphenol oxidase, porcine cytokines, antibody heavy and light chains, lysoxygenase, cytochrome *c* reductase, spinach glycolate oxidase, oncostatin M, lactoferrin, bovine rhodopsin, and transmembrane transferrin. Although expression—or very-high-level expression—of heterologous proteins in *Pichia* has not been successful in every case, there are sufficient examples of success to warrant inclusion of *Pichia* in the standard expression repertoire. In addition to the well-publicized examples of high-level expression (Tschopp et al., 1987a; Clare et al., 1991a; Barr et al., 1992), there are several examples that demonstrate *Pichia*'s behavior as a eukaryote capable of correct post-translational processing, including expression of correctly folded heterologous proteins that contain many disulfide bonds (White et al., 1994), proteins that are active when expressed in *Pichia* but not when expressed in *E. coli* (Yamada et al., 1994; Jenkins and Fayer, 1995; Steinlein

RCUIAS CENTRAL UNIVERSITY LIBRARY

et al., 1995; Vozza et al., 1995), and proteins for which *Pichia* properly recognized the native higher eukaryote secretion signal (Barr et al., 1992; Ridder et al., 1995). Levels of expression vary widely, as do the optimal conditions for expression.

Legal Issues

Heterologous protein expression in *Pichia pastoris* is covered by issued patents that are owned and administered by Research Corporation Technologies (RCT). Invitrogen has the exclusive rights for the distribution of *Pichia* expression technology. The *Pichia* expression system (strains, vectors, and reagents) is available to all researchers simply by purchasing the *Pichia* Expression Kit from Invitrogen. Protein expression in *Pichia*, even for research purposes only, requires a "research license" that is automatically granted with the purchase of the kit from Invitrogen. A "commercial license" is required for expression intended for commercial purposes. The commercial license must be obtained from RCT after the kit is purchased. Questions regarding commercial licenses or the definition of commercial application should be addressed to RCT.

COMMENTARY

Critical Parameters

Successful expression—loosely defined as achieving expression levels that meet the users' needs—can be viewed as a two-step process. The first step is to achieve expression at some level, even if suboptimal. Often successful expression is fully achieved in the first step. If the levels of expression obtained initially are not adequate, however, the second step is optimization of expression. Several parameters can be manipulated to optimize expression levels, including (1) expression as an intracellular protein versus a protein targeted for secretion; (2) expression in a Mut^+ versus a Mut^S recombinant strain; (3) expression in a wild-type versus a protease-deficient mutant; (4) expression in recombinant strains with multiple copies of expression cassettes; and (5) optimization of medium conditions, including specific components (i.e., MM versus BMMY), pH, and expression scale (i.e., flasks versus fermentation). Tools have been developed that allow the researcher to manipulate these parameters and optimize expression. It is essential to keep optimization steps in mind when planning the initial expression experiments.

Other critical tools required for success include good, well-defined methods to unambiguously detect heterologous protein expression. Simple visualization via polyacrylamide gel electrophoresis (PAGE; *UNIT 10.1*) with Coomassie or silver staining (*UNIT 10.5*) is ideal, but other, more sensitive tools—such as antibodies for immunoblots, ELISAs, or immunoprecipitations, or assays for an activity associated with the protein being expressed—are also very useful.

Troubleshooting

Although the *Pichia pastoris* system is easy to use, it requires a number of media that may include components and procedures that are new to some labs. The proper recipes and components should be obtained and the media should be prepared well in advance of beginning the experiments (see *UNIT 5.8*). All the media are very stable and have long shelf lives, so it is feasible to prepare them in advance.

There are several methods to transform *P. pastoris* (Ausubel et al., 1995, Chapter 13), although the spheroplast and electroporation methods are probably the most commonly used. Spheroplast transformation is a relatively straightforward procedure, but it may be the most difficult aspect for first-time users of the system to master. Gentle handling of the cells and meticulous attention to the protocol are required for success.

Most other problems that are likely to arise in expression using *P. pastoris* involve protein expression levels, and are outlined in Table 5.7.5.

Anticipated Results

As the number of expressed proteins in *P. pastoris* increases, more information is being generated to help define routes to successful expression. Based on available data, there is an ~50% to 75% chance of expressing a protein of interest in *P. pastoris* at reasonable levels. The biggest hurdle seems to be generating initial success—that is, expression of the protein at any level. After success at this stage, there are many well-defined parameters that can be manipulated to achieve optimization of expression, and it is often at this stage that acceptable levels of expression are achieved. Although there have been relatively few instances of expression of ≥10 g/liter, there have been many cases of expression in the ≥1 g/liter range, ranking the *P. pastoris* expression system as one of the most productive eukaryotic expression systems available. Likewise, there are several proteins that have been successfully

EMORY UNIVERSITY LIBRARY

Table 5.7.5 Troubleshooting Guide for Failure to Obtain Heterologous Protein

Possible cause of failure	Solution
Expression cassette not present	Demonstrate that recombinant strain contains the gene of interest by PCR or Southern blot hybridization
No transcription	Check for mRNA by RT-PCR or northern blot hybridization; check for AT-rich regions that may be functioning as transcription termination sites within coding region of expressed gene (see Scorer et al., 1993)
No translation of mRNA	Check sequence for possible secondary structure in 5′ untranslated region of mRNA Check codon usage and initiation context for gene of interest
Protein instability	Use protease-deficient strain(s) to prevent proteolysis Change medium or pH medium to optimal for protein
Secretion problem: incomplete secretion, retained in secretory pathway, secreted protein improperly processed, improper modification	Check cell pellet for protein Determine apparent MW of protein
Expression too low	Increase copy number of gene of interest by using pPIC3K, pPIC9K, or pAO815; or scale up fermentation
Expressed protein is deleterious to cells	Examine at earlier time points

expressed in *P. pastoris* following a complete lack of success in baculovirus or *S. cerevisiae*, making the *P. pastoris* system an important alternative in the protein expression toolbox.

Time Considerations

Numerous media are needed for recombinant strain generation and protein expression (Table 5.7.3); medium preparation takes about ½ day and should be completed at least 1 day prior to starting the transformation. The transformation procedure takes ~4 hr with an overnight culture inoculated the day before. It takes 4 to 6 days for prototrophic transformants (e.g., His^+) to form colonies. Screening of transformants to distinguish the His^+ Mut^+ from the His^+ Mut^S takes an additional 3 days. Genomic DNA preps to generate template DNA for PCR analysis and PCR analysis itself can be completed in 1 day. Purification of mRNA for RT-PCR or northern analysis takes ~1 day.

For protein expression experiments in flasks, growth of cultures in glycerol medium to generate biomass takes 1 to 2 days. Methanol induction and a time-course study of protein expression goes from 6 hr to 2 days postinduction for Mut^+ recombinant strains and 1 to 6 days for Mut^S recombinant strains. Analysis of expression by PAGE (*UNIT 10.1*), immunoblotting, ELISA, or specific assay usually requires one additional day.

The total elapsed time from the first transformation to the first protein analysis is ~3 to 4 weeks, with fairly continuous attention required. Commonly, the next phase involves optimization and scale-up once acceptable protein expression levels have been demonstrated. It is reasonable to expect that it will take 6 months from the first transformation to achieving medium- or large-scale production of acceptable levels of heterologous protein expression.

Literature Cited

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) 1995. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Barr, K.A., Hopkins, S.A., and Sreekrishna, K. 1992. Protocol for efficient secretion of HSA developed from *Pichia pastoris*. *Pharm. Eng.* 12:48-51.

Buckholz, R.G. and Gleeson, M.A.G. 1991. Yeast systems for the commercial production of heterologous proteins. *Bio/Technology* 9:1067-1072.

RCU IAS CENTRAL UNIVERSITY LIBRARY

Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. 1991a. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology* 9:455-460.

Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., and Henwood, C.A. 1991b. Production of epidermal growth factor in yeast: High-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* 105:205-212.

Cregg, J.M. and Madden, K.R. 1989. Use of site-specific recombination to regenerate selectable markers. *Mol. Gen. Genet.* 219:320-323.

Cregg, J.M., Barringer, K.J., Hessler, A.Y., and Madden, K.R. 1985. *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* 5:3376-3385.

Cregg, J.M., Tschopp, J.F., Stillman, C., Siegel, R., Akong, M., Craig, W.S., Buckholz, R.G., Madden, K.R., Kellaris, P.A., Davis, G.R., Smiley, B.L., Cruze, J., Torregrossa, R., Velicelebi, G., and Thill, G.P. 1987. High-level expression and efficient assembly of hepatitis b surface antigen in the methylotrophic yeast *Pichia pastoris*. *Bio/Technology* 5:479-485.

Cregg, J.M., Vedvick, T.S., and Raschke, W.C. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Bio/Technology* 11:905-910.

Despreaux, C.W. and Manning, R.F. 1993. The *dacA* gene of *Bacillus stearothermophilus* coding for D-alanine carboxypeptidase: Cloning, structure, and expression in *Escherichia coli* and *Pichia pastoris*. *Gene* 131:35-41.

Digan, M.E., Lair, S.V., Brierley, R.A., Siegel, R.S., Williams, M.E., Ellis, S.B., Kellaris, P.A., Provow, S.A., Craig, W.S., Velicelebi, G., Harpold, M.M., and Thill, G.P. 1989. Continuous production of a novel lysozyme via secretion from the yeast *Pichia pastoris*. *Bio/Technology* 7:160-164.

Ellis, S.B., Brust, P.F., Koutz, P.J., Waters, A.F., Harpold, M.M., and Gingeras, T.R. 1985. Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. *Mol. Cell. Biol.* 5:1111-1121.

Fryxell, K.B., O'Donoghue, K., Graeff, R.M., Lee, H.C., and Branton, W.D. 1995. Fucntional expression of soluble forms of human CD38 in *Escherichia coli* and *Pichia pastoris*. *Protein Express. Purif.* 6:329-336.

Gould, S.J., McCollum, D., Spong, A.P., Heyman, J.A., and Subramani, S. 1992. Development of the yeast *Pichia pastoris* as a model organism for a genetic and molecular analysis of peroxisome assembly. *Yeast* 8:613-628.

Grinna, L.S. and Tschopp, J.F. 1989. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast *Pichia pastoris*. *Yeast* 5:107-115.

Guthrie, C. and Fink, G.R., eds. 1991. Guide to Yeast Genetics and Molecular Biology. *Methods Enzymol.*:194..

Hagenson, M.J., Holden, K.A., Parker, K.A., Wood, P.J., Cruze, J.A., Fuke, M., Hopkins, T.R., and Stroman, D.W. 1989. Expression of streptokinase in *Pichia pastoris* yeast. *Enzyme Microbiol. Technol.* 11:650-656.

Halaas, J.L., Gajiwala, K.S., Maffei, M., Cohen, S.L., Chait, B.T., Rabinowitz, D., Lallone, R.L., Burley, S.K., and Friedman, J.M. 1995. Weight-reducing effects of the plasma protein encoded by the *obese* gene. *Science* 269:543-546.

Howard, P.W. and Maurer, R.A. 1995. A composite Ets/Pit-1 binding site in the prolactin gene can mediate transcriptional responses to multiple signal transduction pathways. *J. Biol. Chem.* 270:20930-20936.

Jenkins, M.C. and Fayer, R. 1995. Cloning and expression of a cDNA encoding an antigenic *Cryptosporidium parvum* protein. *Mol. Biochem. Parasitol.* 71:149-152.

Koutz, P.J., Davis, G.R., Stillman, C., Barringer, K., Cregg, J.M., and Thill, G. 1989. Structural comparison of the *Pichia pastoris* alcohol oxidase genes. *Yeast* 5:167-177.

Larouche, Y., Storme, V., De Meutter, J., Messens, J., and Lauwereys, M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technology* 12:1119-1124.

Ohsawa, I., Hirose, Y., Ishiguro, M., Imai, Y., Ishiura, S., and Kohsaka, S. 1995. Expression, purification, and neurotrophic activity of amyloid precursor protein–secreted forms produced by yeast. *Biochem. Biophys. Res. Commun.* 213:52-58.

Paifer, E., Margolies, E. Cremata, J., Montesino, R., Herrera, L, and Delgado, J.M. 1994. Efficient expression and secretion of recombinant alpha amylase in *Pichia pastoris* using two different signal sequences. *Yeast* 10:1415-1419.

Ratner, M. 1989. Protein expression in yeast. *Bio/Technology* 7:1129-1133.

Reddy., R.G., Yoshimoto, T. Yamamoto, S., and Marnett, L. 1994. Expression, purification, and characterization of porcine leukocyte 12 lipoxygenase produced in methylotrophic yeast, *Pichia pastoris*. *Biochem. Biophys. Res. Commun.* 205:381-388.

Reddy, R., Schmitz, R., Legay, F., and Gram, H. 1995. Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast *Pichia pastoris*. *Bio/Technology* 13:255-260.

Ridder, R., Schmitz, R., Legay, F., and Gram, H. 1995. Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast *Pichia pastoris*. *Bio/Technology* 13:255-260.

Rodriguez, M., Rubiera, R., Penichet, M., Montesinos, R., Cremata, J., Falcon, V., Sanchez, G.,

BOSTON UNIVERSITY MEDICAL LIBRARY

Bringas, R., Cordoves, C., Valdes, M., Lieonart, R., Herrera, L. and de la Fuente, J. 1994. High-level expression of the *b. microplus* Bm86 antigen in the yeast *Pichia pastoris* forming highly immunogenic particles for cattle. *J. Biotechnol.* 33:135-146.

Romanos, M.A., Clare, J.J., Beesley, K.M., Rayment, F.B., Ballantine, S.P., Makoff, A.J., Dougan, G., Fairweather, N.F., and Charles, I.G. 1991. Recombinant *Bordetella pertussis* pertactin (P69) from the yeast *Pichia pastoris*: High-level production and immunological properties. *Vaccine* 9:901-906.

Scorer, C.A., Buckholz, R.G., Clare, J.J., and Romanos, M.A. 1993. The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*. *Gene* 136:111-119.

Scorer, C.A., Clare, J.J., McCombie, W.R., Romanos, M.A., and Sreekrishna, K. 1994. Rapid selection using G418 of high-copy-number transformants of *Pichia pastoris* for high-level foreign gene expression. *Bio/Technology* 12:181-184.

Shiba, K., Ripmaster, T., Suzuki, N., Nichols, R., Plotz, P., Noda, T. and Schimmel, P. 1995. Human alanyl-tRNA synthetase: Conservation in evolution of catalytic core and microhelix recognition. *Biochemistry* 34:10340-10349.

Sreekrishna, K., Nelles, L., Potenz, R., Cruze, J., Mazzaferro, P., Fish, W., Fuke, M., Holden, K., Phelps, D., Wood, P., and Parker, K. 1989. High-level expression purification and characterization of recombinant human tumor necrosis factor synthesized in the methylotrophic yeast *Pichia pastoris*. *Biochemistry* 28:4117-4125.

Steinlein, L.M., Graf, T.N., and Ikeda, R.A. 1995. Production and purification of N-terminal half-transferrin in *Pichia pastoris*. *Protein Express. Purif.* 6-619-624.

Thill, G.P., David, G.R., Stillman, C., Holtz, G., Brierly, R., Engel, M., Buckholz, R., Kenney, J., Provow, S., Vedvick, T., and Siegel, R.S. 1990. Positive and negative effects of multi-copy integrated expression vectors on protein expression in *Pichia pastoris*. *In* Proceedings of the 6th International Symposium on Genetics of Microorganisms, Vol. II (H. Heslot, J. Davies, J. Florent, L. Bobichon, G. Durand, and L. Penasse, eds.) pp. 477-490. Société Française de Microbiologie, Paris.

Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C., and Gingeras, T.R. 1987a. Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucl. Acids Res.* 15:3859-3876.

Tschopp, J.F., Sverlow, G., Kosson, R., Craig, W., and Grinna, L. 1987b. High-level secretion of glycosylated invertase in the methylotrophic yeast *Pichia pastoris*. *Bio/Technology* 5:1305-1308.

Van Nostrand, W.E., Schmaier, A.H., Neiditch, B.R., Siegel, R.S., Raschke, W.C., Sisodia, S.S., and Wagner, S.L. 1994. Expression, purification, and characterization of the Kunitz-type protease inhibitor domain of the amyloid β-protein precursor-like protein-2. *Biochim. Biophys. Acta* 1209:165-170.

Vedvick, T., Buckholz, R.G., Engel, M., Urcan, M., Kinney, J., Provow, S., Siegel, R.S., and Thill, G.P. 1991. High-level secretion of biologically active aprotonin from the yeast *Pichia pastoris*. *J. Ind. Microbiol.* 7:197-201.

Vozza, L., Wittwer, L., Higgins, D.R., Purcell, T.J., Bergseid, M., Collins-Racie, L.A., LaVallie, E.R., and Hoeffler, J.P. 1995. Production of a recombinant bovine enterokinase catalytic subunit in the methylotrophic yeast *Pichia pastoris*. *Bio/Technology*.

Wagner, S.L., Siegel, R.S., Vedvick, T.S., Raschke, W.C., and Van Nostrand, W.E. 1992. High-level expression purification and characterization of the Kunitz-type protease inhibitor domain of protease nixin-2/amyloid β-protein precursor. *Biochem. Biophys. Res. Commun.* 186:1138-1145.

Wegner, G.H. 1990. Emerging applications of the methylotrophic yeasts. *FEMS Microbiol. Rev.* 87:279-283.

Weiss, S., Famulok, M., Edenhofer, F., Wang, Y.H., Jones, I.M., Groschup, M., and Winnacker, E.L. 1995. Overexpression of active Syrian golden hamster prion protein PrP^c as a glutathione-*S*-transferase fusion in heterologous systems. *J. Virol.* 69:4776-4783.

White, C.E., Kempi, N.M., and Komives, E.A. 1994. Expression of higly disulfide-bonded proteins in *Pichia pastoris*. *Structure* 2:1003-1005.

Yamada, M., Azuma, T., Matsuba, T., Iida, H., Suzuki, H., Yamamoto, K., Kohli, Y., and Hori, H. 1994. Secretion of human intracellular aspartic proteinase cathepsin E expressed in the methylotrophic yeast *Pichia pastoris* and characterization of produced recombinant cathepsin E. *Biochim. Biophys. Acta* 1206:279-285.

Key References

Cregg et al., 1993. See above.

Good overview of Pichia pastoris as an expression system with a discussion of glycosylation of expressed proteins.

Scorer et al., 1993. See above.

Describes in detail the advantages of intracellular versus secreted expression, the effect of multiple-copy expression cassettes on expression levels, and an overall approach to optimizing expression levels.

Scorer et al., 1994. See above.

Details the generation of multiple-copy expression cassette recombinant strains using the bacterial Kan^r gene and screening for hyperresistance to G418.

Contributed by David R. Higgins
Invitrogen Corporation
San Diego, California

RCU IASI LIBRARY

Culture of Yeast for the Production of Heterologous Proteins

This unit describes culture of the yeast strains *Saccharomyces cerevisiae* and *Pichia pastoris* for the production of foreign proteins. The theoretical background to foreign gene expression in yeast is discussed in the previous units (*UNITS 5.6 & 5.7*). General techniques for handling yeast are described in Ausubel et al. (1995), Chapter 13, and will not be covered here.

S. cerevisiae expression systems are generally based on the yeast 2μm plasmid, but they differ in choice of promoter and selectable marker, which affect the culture medium and induction conditions that must be used. Vectors with auxotrophic selection markers must be used in appropriate strains (e.g., *LEU2* plasmids in *leu2* strains), and growth must be in appropriate selective medium, i.e., minimal medium containing all the necessary supplements. Vectors with dominant selectable markers (e.g., $G418^r$) can be used in any host strain and selected in rich medium, giving higher growth rate and cell density. The protocols listed here are for three widely used types of promoter: galactose-regulated (*GAL1*, *GAL7*, *GAL10*; Basic Protocol 1), glucose-repressible (e.g., *ADH2*; Alternate Protocol 1), and constitutive glycolytic (e.g., *PGK* or *GAPDH*; Alternate Protocol 2). Minor variations to each can be made depending on the selection system used. The *P. pastoris* expression system uses integrating vectors with the methanol-regulated *AOX1* promoter and *HIS4* selection marker; although transformants are stable, they are generally grown in minimal selective medium.

In most cases, initial analysis is carried out with protein produced in small-scale (5-ml to 1-liter) cultures, and this may be sufficient in cases where the yield is high. Where larger amounts of protein are required, cultures can be grown to high cell density using fermenters. Methods are described for small-scale *S. cerevisiae* (Basic Protocol 1) and *P. pastoris* cultures (Basic Protocol 2) and also for high-density fermentations with these yeasts (Basic Protocols 3 and 4). High-density (>100 g/liter dry weight) growth of *S. cerevisiae* requires computer-controlled carbon source feeding based on continual monitoring of growth parameters (Wang et al., 1979; Fieschko et al., 1987). The protocol given here applies a much simpler feeding strategy based on calculated feed rates (Basic Protocol 3) and yields cell densities of 10 to 30 g/liter (Aiba et al., 1976; Hsieh et al., 1988; Bock Gu et al., 1989). In contrast, with *P. pastoris*, it is simple to obtain extremely high-density cultures (e.g., 130 g/liter) using basic fermenter equipment (Basic Protocol 4). Finally, Support Protocol 1 describes small-scale preparation of protein extracts.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

SMALL-SCALE EXPRESSION USING *S. CEREVISIAE* GALACTOSE-REGULATED VECTORS

BASIC PROTOCOL 1

Galactose-regulated (*GAL*) promoters are induced >1000-fold by addition of galactose but are strongly repressed by glucose, so that in glucose-grown cultures maximal induction can be achieved only following depletion of glucose. For rapid induction, cells can be grown using a nonrepressing carbon source (e.g., raffinose), then induced by direct addition of galactose. First, saturated overnight starter cultures are prepared, and these are used to inoculate fresh medium. These secondary cultures are grown for two doublings (4 to 5 hr including growth lag) to late exponential phase, then induced by direct addition of galactose. Cells are then harvested by centrifugation and may be stored at –70°C as dry pellets or disrupted immediately, using glass beads, to prepare a total protein extract.

Contributed by Michael A. Romanos and Jeffrey J. Clare
Current Protocols in Protein Science (1995) 5.8.1-5.8.17
Copyright © 1997 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY LIBRARY

With either very stable vectors, or vectors having dominant selectable markers, growth and induction can be carried out in rich medium to maximize cell density. Where the plasmid requires auxotrophic selection, minimal selective medium must be used at least for the starter culture, although the induction itself can still usually be carried out in rich medium. The protocol described below is applicable to culture volumes of 5 to 1000 ml.

Materials

YP broth (see recipe) containing 2% (w/v) raffinose *or* YNB medium (see recipe) for auxotrophic selection
50 mg/ml G418 (100×) or supplements for auxotrophic selection, as required
Saccharomyces cerevisiae transformed with a galactose-regulated expression vector and plated on selective medium (see *UNIT 5.6* and Ausubel et al., 1995, Chapter 13)
20% (w/v) galactose (filter sterilize and store at room temperature)
H_2O, sterile

Plastic 25-ml universal containers, sterile (e.g., Nunc)
Heraeus omnifuge centrifuge and 2250 rotor (or equivalent)
15-ml polypropylene snap-top tubes
Centricon concentrator (Amicon; also see *UNIT 4.4*)

1. Place 10 ml YP broth containing 2% raffinose in a sterile 25-ml universal container and inoculate with cells of a single colony of transformed *S. cerevisiae*. Incubate overnight at 30°C with shaking.

   As expression levels may vary between individual transformants, it may be advisable to test several when looking for optimal yield. For vectors with G418-resistance markers, add G418 (0.5 mg/ml final) to the raffinose preinduction medium. For unstable vectors with auxotrophic markers, use YNB medium (see recipe) with 2% raffinose and the appropriate supplements (cell density will be up to 5-fold lower).

   Glucose can be used in place of raffinose in the preinduction growth stages, especially where it is necessary to maintain tight repression of expression, e.g., in expressing toxic proteins. However, excess glucose must be removed by centrifuging and washing the cells in sterile saline prior to resuspending them in induction medium. For glucose-grown cultures, induction will have a several-hour lag period, during which the glucose is exhausted.

2. Measure the A_{600}.

   The A_{600} of an overnight culture should be ~5.0.

3. Dilute the overnight culture into 10 ml YP broth containing 2% raffinose to an A_{600} of 0.25 (typically a 20-fold dilution). Incubate at 30°C with shaking until the culture density reaches an A_{600} of ~1.0 (4 to 5 hr).

4. Add 1 ml of 20% galactose to induce the culture. Incubate 8 to 24 hr (e.g., overnight) with shaking at 30°C.

5. Measure the A_{600} of the culture.

   The A_{600} of the induced cultures should be 10 to 15.

6a. *For intracellular protein expression:* Harvest the cells by centrifuging 5 min at 2000 × *g*, 4°C. Resuspend the cell pellet in 10 ml water, transfer to a 15-ml snap-top tube, and centrifuge again as before. Remove excess water from the pellet. Store at –70°C.

6b. *For secreted protein expression:* Harvest the cells by centrifuging 5 min at 2000 × *g*, 4°C. Collect the medium (supernatant) and concentrate proteins in the medium by

CSU IASI / SENTRAL UNIVERSITY LIBRARY

ultrafiltration using a Centricon concentrator (Amicon). Store the protein concentrate at –70°C. Retain the cell pellet for analysis.

Several secreted proteins have been found to accumulate to lower levels in minimal medium than in rich medium. Yields in minimal medium have been improved by adding Casamino acids (see recipe) or by buffering to pH 6.0 using potassium phosphate (unbuffered media will reach pH ~4).

To concentrate secreted proteins further, concentrate the cells 20-fold by centrifugation prior to induction.

EXPRESSION USING GLUCOSE-REPRESSIBLE *ADH2* VECTORS

ALTERNATE PROTOCOL 1

The *ADH2* promoter is repressed 100-fold by glucose. To maintain repression, cells must always be grown in excess glucose—e.g., 8% (w/v)—because normal concentrations (e.g., 2%) can become depleted. It is also preferable to store transformants on plates containing 8% glucose. Inductions can be carried out by first growing cells in medium containing excess glucose, then inducing by changing to fresh medium containing a nonfermentable carbon source such as raffinose.

Additional Materials (also see Basic Protocol 1)

YP broth (see recipe) containing 8% (w/v) glucose
Saccharomyces cerevisiae transformed with a glucose-repressible *ADH2* expression vector and plated on selective medium (see UNIT 5.6 and Ausubel et al., 1995, Chapter 13)
YP broth (see recipe) containing either 2% (w/v) raffinose and 2% (v/v) ethanol *or* 1% (w/v) glucose

100-ml Erlenmeyer flask, sterile
Diastix glucose test strips (Bayer Diagnostics; optional)

1. Prepare a 10-ml exponential phase culture of the *S. cerevisiae* transformant in YP broth containing 8% glucose (see Basic Protocol 1, steps 1 to 3).

   Depending on the vector characteristics, selective medium may need to be used, as in Basic Protocol 1.

2. Harvest the cells by centrifuging 5 min at 2000 × *g*, 20°C. Resuspend the pellet in 10 ml sterile water.

3. Centrifuge again as in step 3 and resuspend the pellet in 10 ml YP broth containing 2% raffinose/2% ethanol.

4. Incubate 8 to 24 hr (e.g., overnight) at 30°C under conditions of maximum aeration (e.g., by incubating in a 100-ml Erlenmeyer flask with vigorous agitation).

   Induction starts after a lag of several hours.

   As an alternative to steps 3 and 4, use the overnight starter culture in 8% glucose to inoculate YP broth containing 1% glucose to a starting A_{600} of 0.05 (typically a 100-fold dilution). Incubate 18 to 24 hr at 30°C with vigorous aeration. Monitor glucose depletion using Diastix glucose test strips if possible.

5. Measure the A_{600}.

6. Harvest the cells by centrifuging 5 min at 2000 × *g*, 4°C.

7. Analyze the cell pellet for proteins expressed intracellularly and analyze the cell pellet and the medium (supernatant) for secreted proteins.

BOULDER / CENTRAL UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 2

EXPRESSION USING VECTORS WITH GLYCOLYTIC GENE PROMOTERS

A number of commonly used expression vectors have constitutive promoters from glycolytic genes (e.g., *PGK*, *GAPDH*, and *ADH1*). Because instability problems are more frequent with constitutive vectors, it is recommended that selective media be used throughout.

Materials (also see Basic Protocol 1)

YNB medium (see recipe) containing 2% (w/v) glucose and the required supplements

Saccharomyces cerevisiae transformed with an expression vector containing a glycolytic gene promotor and plated on selective medium (see *UNIT 5.6* and Ausubel et al., 1995, Chapter 13)

1. Prepare an overnight starter culture of the *S. cerevisiae* transformant in YNB medium containing 2% glucose and required supplements (see Basic Protocol, step 1).

   Doubling time for the transformed strain will usually be longer than for the untransformed strain. Constitutive high-level expression using vectors that contain glycolytic gene promotors causes a metabolic burden on the cell that usually results in reduced growth rates. The degree of the effect will depend on the gene expressed, but doubling times may increase to 3 or 4 hr, or greater in extreme cases. Very long doubling times are diagnostic of toxicity, and indicate that a regulated vector should be used.

2. Measure the A_{600}.

   The A_{600} of the overnight culture should be 5.

3. Dilute the overnight culture into 10 ml fresh supplemented YNB medium to an A_{600} of 0.0025.

   Using medium with a higher glucose concentration—e.g., 8% (w/v)—may improve expression levels, because glycolytic promoters are partly glucose-inducible.

4. Incubate at 30°C with shaking until the A_{600} reaches 0.25 to 0.5—18 to 24 hr, depending on the foreign protein.

5. Harvest the cells for analysis by centrifuging 5 min at 2000 × *g*, 4°C.

6. Analyze the cell pellet for proteins expressed intracellularly and analyze the cell pellet and the medium (supernatant) for secreted proteins.

BASIC PROTOCOL 2

SMALL-SCALE EXPRESSION IN *PICHIA PASTORIS*

Small-scale cultures of *Pichia pastoris* are used for the initial analysis of expressed proteins and for monitoring established expression strains. As different transformants may vary widely in product yield, it may be necessary to screen a significant number (see Critical Parameters and *UNIT 5.7*). Aeration is a key consideration because the efficiency of induction is reduced when oxygen is limited. To avoid this, incubate with vigorous shaking, keep the culture volume to a minimum (e.g., 5 to 10 ml) and use large flasks (e.g., >100 ml, preferably baffled) with loose or permeable covers. Some evaporation may occur, and during extended inductions (>1 day) it may be necessary to add further aliquots of methanol. The optimal induction period may vary for different proteins, and it is recommended that a time course be carried out to maximize yields.

CSU IAS CENTRAL UNIVERSITY LIBRARY

Materials

YNB medium (see recipe) containing 2% (w/v) glycerol
Pichia pastoris transformed with an expression vector and plated on selective medium (see *UNIT 5.7* and Ausubel et al., 1995, Chapter 13)
0.002 mg/ml biotin (500×; filter sterilize and store <1 year at 4°C)
H_2O, sterile
YNB medium (see recipe) containing 1% (v/v) methanol

Plastic 25-ml universal containers, sterile
Heraeus omnifuge centrifuge and 2250 rotor (or equivalent)
100-ml Erlenmeyer flasks, sterile
15-ml polypropylene snap-top tubes

1. Place 10 ml YNB broth containing 2% glycerol in a sterile 25-ml universal container and inoculate with cells of a single colony of transformed *P. pastoris*. Incubate overnight at 30°C with shaking.

2. Measure the A_{600}.

   The A_{600} of the overnight culture should be ~5.

3. Dilute the overnight culture into 10 ml YNB medium containing 2% glycerol to an A_{600} of 0.25. Incubate 6 to 8 hr at 30°C with vigorous shaking to obtain an exponentially growing culture.

4. Collect the cells by centrifuging 5 min at 2000 × *g*, 4°C. Resuspend the cell pellet in 10 ml sterile water.

5. Pellet the cells by centrifuging 5 min at 2000 × *g*, 4°C. Resuspend the cell pellet in 10 ml YNB medium containing 1% (v/v) methanol. Incubate 1 to 5 days at 30°C with maximum aeration (e.g., incubate culture in a 100-ml Erlenmeyer flask with vigorous shaking).

   If inducing >1 day, add an additional 0.5% methanol on the second day. The final density of the culture should be A_{600} = 5 to 10. The concentration of a secreted product in the medium is generally proportional to the density of the culture. Thus, to increase the concentration of a secreted protein, grow a larger culture (e.g., 100 ml) and concentrate the cells prior to the induction by resuspending in 5 ml of the methanol-containing induction medium. In addition, the yield of some secreted proteins may be reduced by extracellular proteases. This can sometimes be ameliorated by altering the composition of the medium, for example by buffering the pH or by adding amino acid–rich supplements such as Casamino acids (see recipe) or peptone (Clare et al., 1991b; Cregg et al., 1993).

6. Collect the cells by centrifuging 5 min at 2000 × *g*, 4°C. Resuspend the cell pellet in 10 ml water. If the expressed protein is secreted, retain the supernatant, concentrating the protein by ultrafiltration (see Basic Protocol 1, step 6b).

7. Transfer the cell suspension to a 15-ml polypropylene snap-top tube. Centrifuge 5 min at 2000 × *g*, 4°C. Remove excess water from the pellet and store the washed cell pellet at –70°C.

8. Analyze the cell pellet for intracellularly expressed protein and analyze the cell pellet and the supernatant for secreted protein.

BCU/ASTON/CLARE UNIVERSITY LIBRARY

LARGE-SCALE EXPRESSION USING *S. CEREVISIAE* GALACTOSE-REGULATED VECTORS

High-density fermentation of *S. cerevisiae* is problematic because the yeast has a limited capacity to metabolize glucose oxidatively; ethanol accumulates, which eventually inhibits its growth to a specific growth rate of 0.2 to 0.25 hr^{-1} (Sonnlietner and Kappeli, 1986). Additionally, glucose concentrations above ~0.1% repress respiration. To prevent the conversion of glucose to ethanol and to maximize biomass, *S. cerevisiae* is grown using a fed-batch process: glucose is slowly fed to the fermenter, the feed rate increasing as the cell density increases, so as to obtain an almost constant growth rate (μ). The growth rate above which ethanol is formed (μ_{crit}), and hence the critical feed rate, are determined by both the strain and the growth medium. Figure 5.8.1 shows the relationship between the growth rate and biomass yield, and the effect of exceeding the μ_{crit} (Pirt, 1985). The following simplified protocol uses calculated feed rates.

Materials

YNB medium (see recipe) containing 1% (w/v) glucose and appropriate supplements
Saccharomyces cerevisiae transformed with a galactose-regulated expression vector and plated on selective medium (see *UNIT 5.6* and Ausubel et al., 1995, Chapter 13)
BSSM fermenter medium (see recipe)
20% (w/v) glucose
20% (w/v) galactose
3 M phosphoric acid
3 M sodium hydroxide
Antifoam (Dow Corning M30 antifoam)

250-ml and 2-liter Erlenmeyer flasks, sterile
Benchtop fermenter, 5-liter total volume, with data logging and MENTOR control system (Life Science Laboratories)
Electronic balance (Mettler) with RS232 communications
Sorvall RC-3B centrifuge and H6000A rotor (or equivalent)

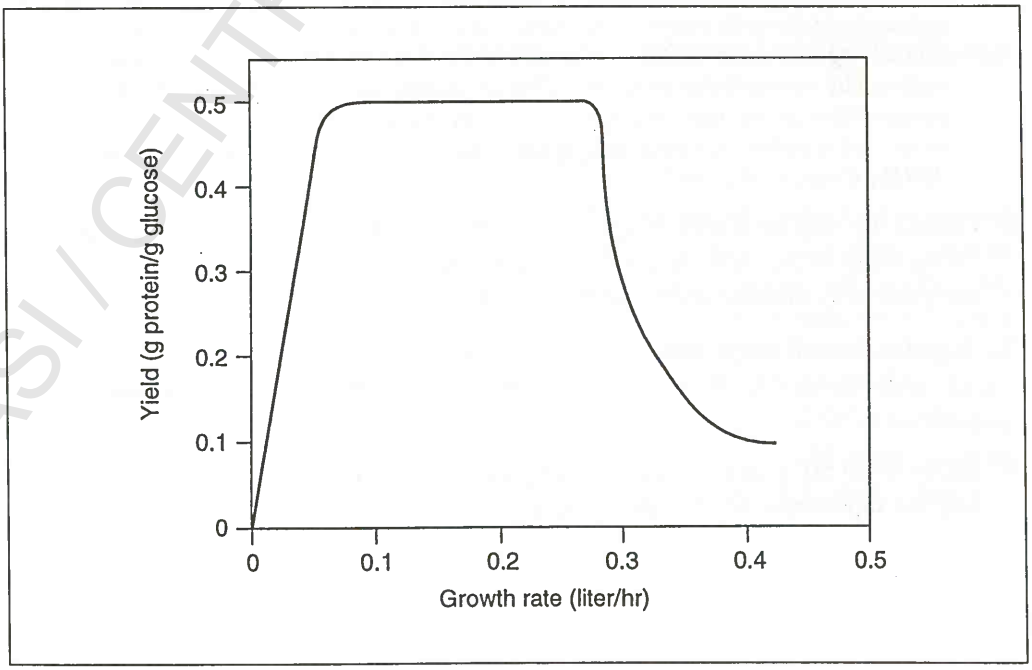


Figure 5.8.1 Effect of growth rate on yield.

EGU IASI / CENTRAL UNIVERSITY LIBRARY

Prepare the cultures

1. Place 50 ml YNB medium containing 1% glucose and supplements in a sterile 250-ml Erlenmeyer flask and inoculate with cells from a single colony of transformed *S. cerevisiae*. Incubate overnight at 30°C with shaking.
2. Measure the A_{600}.

   The A_{600} of the overnight culture should be ~4.0.
3. Dilute 10 ml of the overnight culture into 250 ml YNB medium containing 1% glucose in a 2-liter shaker flask. Incubate 18 to 24 hr at 30°C with shaking (150 rpm).
4. Measure the A_{600}.

   The A_{600} should be ~4.0.

Calculate the feed rate

5. Determine the feed rate for the fermenter using the following equation (see Table 5.8.1 for definitions of terms) for a 20% glucose solution, where $C1 = (X_0 \times V \times \rho)/(Y \times Sf)$.

   $$\text{cumulative weight of feed} = (C1 \times e^{\mu t}) - C1$$

   Based on the approximate values for a 5-liter fermentation protocol given in Table 5.8.1, C1 = 2.7, and the feeding algorithm reduces to:

   $$\text{cumulative weight of feed} = (2.7 \times e^{0.16t}) - 2.7$$
6. Determine the feed profile by calculating the cumulative weight of feed for $t = 1, 2, \ldots, n$ hr. Enter these values into the MENTOR computer system.

Express the protein

7. Place the fermenter reservoir containing 20% glucose feed on an electronic balance linked to MENTOR, which will control the feed rate via a dosing pump.

   The fermentation consists of two phases: a glucose fed-batch phase followed by batch addition of galactose to induce expression.

Table 5.8.1 Variables for Calculating *S. cerevisiae* Feeding Algorithm

Variable	Definition	Value for 5-liter protocol	Comment
X_o	Biomass in fermenter at start of feeding (g/liter)	0.07	Derived by multiplying the starting A_{600} in the fermenter by 0.22 (assume shaker flask inoculum has $A_{600} = 4$)
V	Working volume of fermenter (liter)	3.5	Maximum working volume
Y	Yield coefficient (g cells per g glucose)	0.45	Theoretical value is 0.5, but a more conservative value of 0.45 is recommended
Sf	Glucose concentration in feedstock (g/liter)	200	—
μ	Target growth rate (hr^{-1})	0.16	Suggested value (see text for optimization)
ρ	Feed density (g/liter)	1000	Approximation is sufficient

BCUIAS / CENTRAL UNIVERSITY LIBRARY

8. Add 2.5 liters BSSM fermentation medium to the fermenter. Then add the culture from step 2 to the fermenter and initiate feeding with 20% glucose. Feed for 28 hr.

Monitor the A_{600} closely during feeding; if the growth rate drops below the target ($\lambda = 0.16\ hr^{-1}$), stop glucose feeding early. The growth rate selected in this protocol serves as a useful starting point for any strain. To optimize the protocol, determine the maximum growth rate (μ_{max}) for the strain in shaker-flask cultures, then calculate the feed rate, setting the desired growth rate at $\lambda = 0.5\mu_{max}$. This should be safely below μ_{crit}; if necessary, increase λ from this value. If the feed rate is too high, glucose or ethanol will be detected in the fermentation and biomass yield will be reduced. The effect of altering λ on the feeding profile is shown in Figure 5.8.2.

9. Induce the culture by adding 250 ml of 20% galactose. Incubate the culture 24 hr.

If glucose has accumulated, then induction will be delayed for several hours.

Higher expression levels have been reported when 1% (w/v) galactose is pulsed in during the glucose feeding stage, and the glucose feed is then replaced with a galactose feed (Fieschko et al. 1987).

The duration of induction should be optimized by determining the kinetics of product formation. Harvesting shortly before the culture stops growing may help minimize proteolytic degradation of the product. Cells in stationary phase are more resistant to mechanical breakage.

10. Harvest the culture after induction by centrifuging 10 min at 7000 × *g* in a Sorvall RC-3B centrifuge with H6000A rotor.

11. Analyze the cell pellet for intracellularly expressed protein and analyze the cell pellet and the supernatant for secreted protein.

BASIC PROTOCOL 4

LARGE-SCALE EXPRESSION USING *PICHIA PASTORIS*

Optimal expression in *P. pastoris* requires the use of fermenters, which increases both the efficiency of induction and the culture density. The method is simple, requiring only a basic fermenter with monitors and controls for pH, dissolved O_2, stirring speed, temperature, and air flow. Cultures are grown in a defined medium containing glycerol. The pH is maintained by the addition of ammonium hydroxide, which also serves as the nitrogen source. Dissolved O_2 must be maintained at >20% for optimal growth. Fermentations are carried out in three stages: an initial batch growth phase, a glycerol-limited phase during which the culture reaches high cell density, and, finally, a methanol-induction phase of

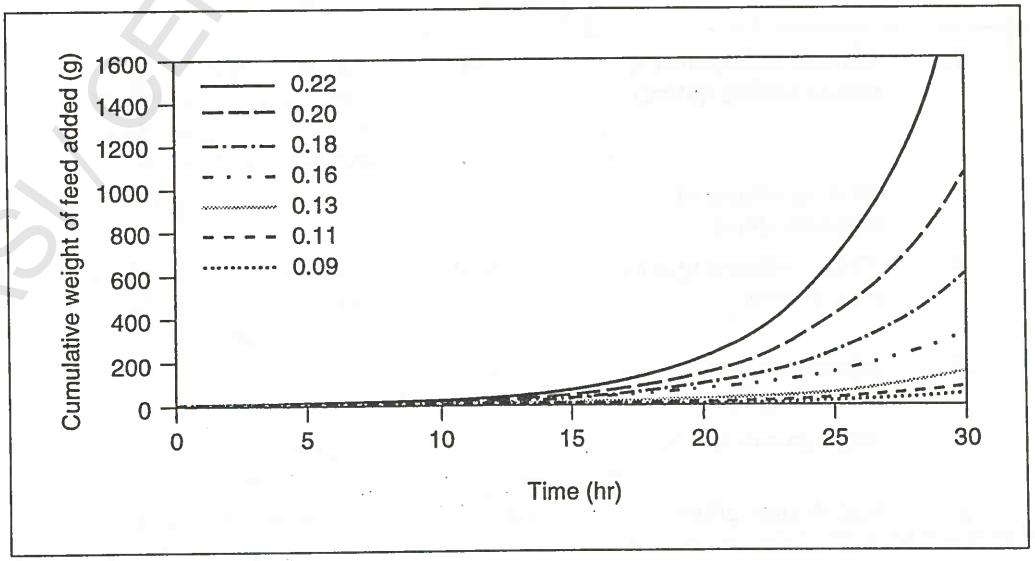


Figure 5.8.2 Effect of growth rate on feed profile.

PCU IASL CENTRAL UNIVERSITY LIBRARY

several days during which product accumulates. The protocol given below is for a 1-liter "bench-top" fermenter, but very similar conditions have been used for cultures up to 240 liters (Cregg et al., 1987).

Materials

YNB medium (see recipe) containing 2% (v/v) glycerol
Pichia pastoris transformed with an expression vector and plated on selective medium (see *UNIT 5.7* and Ausubel et al., 1995, Chapter 13)
5× basal salts solution (see recipe)
PTM_1 solution (see recipe)
Glycerol/PTM_1: 12 ml PTM_1 solution/liter 50% (v/v) glycerol
50% (v/v) ammonium hydroxide
Methanol/PTM_1: 12 ml PTM_1 solution/liter 100% methanol

Fermenter equipped with monitors or controls for pH, temperature, dissolved O_2, airflow, and stirring speed
Erlenmeyer flask, sterile

1. Place 10 to 50 ml YNB medium containing 2% glycerol in an Erlenmeyer flask and inoculate with cells of a single colony of transformed *P. pastoris*. Incubate overnight at 30°C with shaking.
2. Add 1 liter of 5× basal salt solution, 4 ml PTM_1 solution, and 5% (v/v) glycerol to the fermenter and mix well.
3. Add the overnight culture to the fermenter. Incubate at 30°C until the glycerol is exhausted (~24 to 30 hr). Adjust the aeration and agitation to maintain dissolved O_2 at >20%, and maintain the pH at 5.0 with 50% ammonium hydroxide.

   The dissolved O_2 level will begin to rise sharply when the glycerol becomes exhausted.
4. Initiate a limited glycerol/PTM_1 feed at 12 ml/hr and continue for 4 to 24 hr, maintaining pH, temperature, and dissolved O_2.

   During this period the culture should grow to a density of 30 to 100 g dry weight/liter. Cell density can be monitored by measuring A_{600}; 1 A_{600} unit = ~0.33 g/liter, although different spectrophotometers will vary and each should be calibrated individually.

   To achieve optimal yield, the duration of this glycerol-limited phase, and thus the cell density at induction, can be varied depending on the type of product and the Mut phenotype of the host (see UNIT 5.7). Mut^+ strains can be induced at low cell density because they continue to grow during induction. Mut^S strains should be allowed to achieve high density before induction. The yield of secreted proteins may be improved by early induction, allowing secretion to occur during a period of active growth.
5. Induce the culture by replacing the glycerol feed with a methanol/PTM_1 feed at the rate of 1 ml/hr. Gradually increase the methanol feed rate over a period of 6 hr to 6 to 10 ml/hr and continue fermentation for an additional 46 to 92 hr.

   For efficient induction of Mut^+ strains, the aim is to achieve maximal methanol feed rates (after an initial period to adapt the culture to methanol utilization) while maintaining sufficient aeration. If dissolved oxygen levels cannot be maintained >20%, the methanol feed rate should be reduced. This will also prevent accumulation of toxic levels of methanol. To avoid this problem with Mut^S strains, the methanol feed rate should be reduced to 1 ml/hr after 48 hr of induction.

B. C. U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI - CENTRAL UNIVERSITY LIBRARY

SUPPORT PROTOCOL

SMALL-SCALE PREPARATION OF PROTEIN EXTRACTS

Following induction using one of the previous basic or alternate protocols, protein extracts may be made from the harvested yeast cells either by mechanical disruption or by spheroplasting followed by lysis in hypotonic buffer. A simple mechanical disruption method is described here, in which cells are vortexed in the presence of glass beads. This method can be used for both analytical- (e.g., 10-ml) or preparative-scale cultures; for the latter a Bead Beater (BioSpec Products) may be used. Cells from 10-ml inductions can be disrupted rapidly using glass beads on a vortex mixer with tube rack. For *Pichia pastoris* high-density fermentations of >2 liters, a larger bead mill with efficient cooling is required (Dyno Mill). Although proteases are less of a problem than with *Escherichia coli*, cell breakage should be carried out in the cold, ideally in the presence of protease inhibitors. A typical extraction buffer is given here, but the particular buffer used will depend on the protein being studied and the subsequent analysis or purification to be performed. If the aim is simply to detect a polypeptide by SDS-polyacrylamide gel electrophoresis (*UNIT 10.1*), the composition of the buffer is less important, because the total extract is denatured before analysis.

Materials

Cell pellet from an induced culture of transformed yeast cells
Protein extraction buffer (see recipe), ice-cold
5× protease inhibitor cocktail (see recipe)

15-ml polypropylene snap-top tubes
0.45-mm glass beads, acid washed (see recipe)
Platform vortex mixer (e.g., IKA-Vibrax-VXR with test tube rack)

1. Resuspend the cell pellet from an induced culture of transformed yeast cells in 0.5 ml ice-cold extraction buffer plus 125 µl of 5× protease inhibitor cocktail in a 15-ml polypropylene snap-cap tube.
2. Add glass beads to two-thirds the height of the meniscus. Place the tubes on a platform vortex mixer in a cold room and vortex 10 min at full speed.

   Samples can be vortexed by hand using a vortex mixer in six 30-sec bursts punctuated by 30-sec intervals on ice. Breakage can be checked using a phase-contrast light microscope: intact cells appear bright and broken cells are phase-dark "ghosts." Breakage is most efficient in wide round-bottom tubes. For larger volumes use 30-ml centrifuge tubes or a Bead Beater.
3. Recover the cell extract from the glass beads using a fine-tipped plastic Pasteur pipette. To recover as much of the extract as possible, reextract the beads with 0.5 ml ice-cold extraction buffer.

   Alternatively, centrifuge the cell extract and beads 2 min at 1000 × g, 4°C, through a plastic disposable filter (e.g., Poly-Prep chromatography columns, Bio-Rad). Check that the protein of interest is not retained by the filter.

   For large-scale purification of cytosolic proteins, the preparation may be centrifuged 1 hr at 100,000 × g, 4°C, to pellet membranes.
4. Determine the total protein concentration of the extract using a standard assay (e.g., Bio-Rad protein assay; also see *UNIT 3.4*).

   The yield of total protein from a 10-ml culture grown to A_{600} = 10 is ~10 mg.

PCU IASI / FEDERAL UNIVERSITY LIBRARY

5. If desired, separate insoluble protein by microcentrifuging 15 min at 12,000 × *g*, 4°C. Transfer the soluble protein to a new tube and determine the protein concentration; retain this and the pellet containing insoluble proteins for analysis.

   To determine the proportion of a protein that is soluble, resuspend the insoluble protein pellet in the same volume of buffer as the soluble protein preparation, and analyze equal volumes of total, soluble, and insoluble protein by SDS-PAGE (see UNIT 10.1).

6. Store protein extracts at –20° or –70°C.

   Upon thawing, a precipitate, largely consisting of cell membranes, may appear; this should be dispersed prior to SDS-PAGE (e.g., by pipetting up and down).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Basal salts solution, 5×

Per liter:
42 ml phosphoric acid (0.62 M final)
1.8 g calcium sulfate·$2H_2O$ (10 mM final)
28.6 g potassium sulfate (0.16 M final)
23.4 g magnesium sulfate·$7H_2O$ (10 mM final)
6.5 g potassium hydroxide (0.12 M final)
Autoclave
Store indefinitely at room temperature

BSSM fermenter medium

3.35 g monobasic potassium phosphate (1 mM final)
1.8 g magnesium sulfate·$7H_2O$ (3 mM final)
26.25 g ammonium sulfate (0.08 M final)
2.5 liters distilled water
0.6 ml 0.07 M manganese sulfate·$4H_2O$ (20 µM final)
0.6 ml 0.08 M copper sulfate·$5H_2O$ (20 µM final)
0.6 ml 0.6 mM zinc sulfate·$7H_2O$ (4 µM final)
0.1 ml 33 mM potassium iodide (1 µM final)
0.1 ml 0.39 M sodium molybdate·$2H_2O$ (20 µM final)
0.1 ml 0.08 M ferric chloride·$6H_2O$ (3 µM final)
Place in 5-liter fermenter vessel and autoclave
Add 50 ml vitamin solution (see recipe)

Casamino acids, 10×

Prepare a 20% (w/v) solution of Casamino acids (low-salt, Difco). Autoclave. Store up to 1 year at 4°C.

Glass beads, acid washed

Wash 0.45-mm-diameter glass beads (B. Braun) for 16 hr in concentrated HCl. Rinse thoroughly in distilled water. Bake 16 hr at >150°C. Store indefinitely at room temperature.

BCU FISH CENTRAL UNIVERSITY LIBRARY

Protease inhibitor cocktail, 5×

20 mM EDTA
20 mM EGTA
20 mM PMSF
10 mg/ml pepstatin
10 mg/ml leupeptin
10 mg/ml chymostatin
10 mg/ml antipain
Store in 1-ml aliquots up to 1 year at –20°C

Protein extraction buffer

50 mM Tris·Cl, pH 7.4
100 mM NaCl
2 mM EDTA
Autoclave
Store indefinitely at room temperature
Chill on ice before use

PTM_1 salt solution

Per liter:
6 g cupric sulfate·$5H_2O$ (24 mM final)
0.08 g potassium iodide (0.5 mM final)
3 g manganese sulfate·H_2O (18 mM final)
0.2 g sodium molybdate·$2H_2O$ (0.8 mM final)
0.02 g boric acid (0.3 mM final)
0.5 g cobalt chloride·$6H_2O$ (2 mM final)
20 g zinc chloride (0.15 M final)
65 g ferrous sulfate·$7H_2O$ (0.23 M final)
0.2 g biotin (0.8 mM final)
5 ml sulfuric acid (0.9 M final)
Autoclave
Store indefinitely at room temperature

Vitamin solution

Per 50 ml:
10 mg biotin
65 mg pantothenic acid
100 mg *meso*-inositol
5 mg 4-aminobenzoic acid
65 mg pyridoxine·HCl
50 mg thiamine
50 mg nicotinic acid
Filter sterilize
Store up to 1 year at 4°C

The vitamin solution can also be prepared as a 10× stock.

YNB stock, 10×

Dissolve 6.7 g yeast nitrogen base without amino acids (Difco) in 100 ml distilled water. Filter sterilize with 0.2-μm filter. Store up to 1 year at 4°C.

BCU IASI CENTRAL UNIVERSITY LIBRARY

YNB medium

Add 10× YNB stock (see recipe) and 500× biotin stock to sterile distilled water to give 1× final concentration for basic YNB medium. Add a carbon source (e.g., glucose or raffinose) and nutritional supplements (see Table 5.8.2) according to vector selection requirements. Store up to 1 year at 4°C.

Table 5.8.2 Nutritional Supplements for YNB Medium

Nutrient[a]	Final concentration[b] (mg/liter)
Adenine sulphate	20
L-Arginine·HCl	20
L-Aspartic acid	100
L-Glutamic acid	100
L-Histidine·HCl	20
L-Isoleucine	30
L-Leucine	30
L-Lysine·HCl	30
L-Methionine	20
L-Phenylalanine	50
L-Serine	375
L-Threonine	200
L-Tryptophan	20
L-Tyrosine	30
Uracil	20
Valine	150

[a]Prepare as 100× stock solutions and filter sterilize. Store at 4°C.
[b]Add to filter-sterilized medium.

YP broth

10 g yeast extract
20 g Bacto Peptone
1 liter H_2O
Autoclave
Store up to 1 year at room temperature

COMMENTARY

Background Information

Yeast expression systems have advantages over other commonly used systems for the production of proteins in certain situations (see *UNITS 5.6 & 5.7* and Romanos et al., 1992). Although typical percentage yields of proteins in *Saccharomyces cerevisiae* (1% to 5% of cell protein) are less than those obtained with *Escherichia coli* or baculovirus, the volumetric yield with *S. cerevisiae* may be higher than with the baculovirus system due to the much higher cell density. Depending on the productivity of the strain, shaker-flask inductions, scaled up to 1 liter if necessary, may be sufficient to provide the amount of protein needed. Where a significant amount of protein is required, e.g., >100 mg, fermenters may be used to culture yeast to extremely high density, e.g., 100 g dry weight/liter; this is particularly valuable with secretion, where the concentration of protein in the medium can increase almost proportionately. Because of the technical complexity involved in achieving such densities with *S. cerevisiae*, this unit includes a much simplified protocol using calculated feeding rates, which gives lower cell densities. However, the au-

COLUMBIA CENTRAL UNIVERSITY LIBRARY

Table 5.8.3 Comparison of Yields for Heterologous Expression of Two Proteins

Expression system	Tetanus toxin fragment C (% of total protein)	Pertactin (% of total protein)
E. coli	24	30
Baculovirus	10[a]	>40[a]
S. cerevisiae GAL7	2-3	0.1
P. pastoris	28[b]	10[c]

[a]Estimate; variable yield greatly reduced on scale-up.
[b]Scale-up without loss of yield to >12 g/liter.
[c]Scale-up without loss of yield to >4 g/liter.

thors recommend the use of *Pichia pastoris* for high-density fermentation wherever possible.

P. pastoris is an industrial yeast, initially selected for its high-density growth in simple defined media; this organism has also been developed as an expression system (Romanos et al., 1992; Cregg et al., 1993). There are numerous cases in which much higher percentage yields of intracellular proteins have been obtained using *P. pastoris* than using *S. cerevisiae*, often to levels equivalent to those in *E. coli* and baculovirus—e.g., 5% to 40% of cell protein (Sreekrishna et al., 1989; Clare et al., 1991a; Romanos et al., 1991). Table 5.8.3 shows comparative data for two proteins expressed in *E. coli*, baculovirus, *S. cerevisiae*, and *P. pastoris*. *P. pastoris* is also generally more efficient in secreting foreign proteins (Clare et al., 1991b). Moreover, scale-up of *P. pastoris* to high-density fermentations is very simple and can give enormous yields on a volumetric basis. The tight regulation of the *AOX1* promoter means that the growth phase is essentially identical for different protein products, and scale-up to large fermenters does not affect biomass or product yield. Additionally, contrary to the experience with *S. cerevisiae*, an immediate improvement in percentage yield of foreign protein is generally seen on going from shaker flasks to fermenters. This appears to be due at least in part to the high O_2 demand of the organism, which is not satisfied in shaker flasks. For further general information on use of the *P. pastoris* system, consult the instruction manual available from Invitrogen and UNIT 5.7.

Critical Parameters

S. cerevisiae shaker-flask inductions are relatively straightforward, but improvements can be made by testing different host strains, including protease-deficient strains. In *S. cerevisiae* fed-batch fermentations, achieving the correct feed rate is critical; overfeeding, especially in the first few hours, will result in lower cell yield. Because the feed rate in the protocol given here is calculated, it is essential that the density of the inoculum reaches the required level. Another critical factor is dissolved oxygen: levels <30% of saturation affect cell growth, while <10% may affect expression. Increasing the agitation or air-flow rate can increase dissolved oxygen.

The most important factors affecting expression in *P. pastoris* have been discussed in reviews (e.g., Romanos et al., 1992; Cregg et al., 1993; UNIT 5.7). When using *P. pastoris*, selection of the transformant and growth conditions are highly critical. Selection of the transformant is important because *P. pastoris* transformations, especially *Bgl*II transplacements, result in a highly heterogeneous population of transformants. The variables are Mut phenotype (determined by whether or not *AOX1* is disrupted), site of integration, and copy number. For intracellular proteins this usually results in extreme clonal variation in yield between different transformants, dependent mainly on the number of copies of the vector integrated into the chromosome (Sreekrishna et al., 1989; Clare et al., 1991a; Romanos et al., 1991; Scorer et al., 1994). A number of approaches have been used to maximize vector copy number—e.g., screening transformants by DNA dot blot (Romanos et al., 1991) and G418 selection (Scorer et al., 1994). Although it appears that the highest copy numbers result from transplacements using the spheroplast technique, for rapid isolation of near-optimal transformants the authors recommend the much simpler method of G418 selection following electroporation (Scorer et al., 1994).

With secretion of foreign proteins, by contrast, the presence of many copies is sometimes

EDUCATIONAL UNIVERSITY LIBRARY

suboptimal, and the Mut phenotype may have a greater influence on yields; consequently, a significant number of transformants with different copy numbers may need to be tested for maximum yield (Thill et al., 1990; Clare et al., 1991b). An alternative approach is to preselect efficient secretors empirically by micro-scale inductions in 96-well plates (Laroche et al., 1994).

A critical factor during *P. pastoris* growth is its tendency to become oxygen-limited in small-scale inductions, leading to low and variable yields. This effect can be minimized by taking steps to ensure good aeration. However, it is nevertheless a major reason for switching to fermenters as soon as the most efficiently expressing transformants have been identified from small-scale inductions. The change from small-scale to fermenter is usually accompanied by an increase in percentage yield.

A factor that affects *P. pastoris* fermentations is the Mut phenotype, which determines whether or not there is continued growth during the methanol induction phase. Mut^+ strains have a greater tendency to become oxygen-limited during induction and may give more variable yields in small-scale inductions. The methanol feeding rate will be different for Mut^+ and Mut^S strains because the latter can be poisoned by excess methanol feeding.

Troubleshooting

Problems in expression can arise at numerous stages from vector stability through to protein turnover. The following problems are among the most common: (1) genetic instability of the transformed yeast strain, particularly on scale-up; (2) inability to produce some proteins because of their high toxicity; (3) failure to express AT-rich genes due to premature transcriptional termination; (4) inefficient secretion of some larger (>30 kDa) proteins; (5) proteolysis of secreted proteins; and (6) inappropriate glycosylation of mammalian glycoproteins. These problems have been discussed in previous reviews (e.g., Romanos et al., 1992; *UNITS 5.6 & 5.7*), and this section will concentrate on those problems related purely to strain stability and growth conditions.

In *S. cerevisiae*, low yields of foreign proteins may arise because of reduced copy number, rearrangement, or mutation of the vector, most frequently during the constitutive expression of toxic proteins. The solution to most of these problems is simply to use a tightly regulated promoter—e.g., a *GAL* promoter or the *P. pastoris AOX1* promoter—to reduce selection pressure during the growth phase of the culture. It is also good practice to always prepare banks of −70°C glycerol stocks of selected transformants before the start of a project, so that the starting point of each culture is the same. However, extreme product toxicity can be an intractable problem, and it may then be more profitable to try a transient expression system, such as baculovirus.

In *P. pastoris*, a common problem is proteolysis of secreted proteins (reviewed in Cregg et al., 1993). The problem may occur in shaker flasks but is exacerbated in high-density fermenter cultures, possibly due to the higher concentration of proteases and the different medium used. A number of strategies have been used successfully to overcome proteolysis: adding peptide supplements (peptone or Casamino acids), buffering the medium to pH 3 or 6, or using a *pep4* protease-deficient strain.

Anticipated Results

Yields of intracellular proteins from *S. cerevisiae* are usually in the range of 0.5% to 5% of total cell protein, or 1.5 to 15 mg/liter/A_{600} unit (15 to 150 mg/liter for typical culture densities in YP medium). Many small non-glycosylated polypeptides are efficiently secreted and may reach levels of 10 mg/liter/A_{600} unit in the culture supernatant. Larger proteins (>25 kDa) are secreted far less predictably, some efficiently and some poorly. The cell density at the end of the fed-batch phase of *S. cerevisiae* fermentations should reach at least $A_{600} = 20$ (4.0 g/liter) and $A_{600} = 40$ (8 g/liter) at harvest. Using this protocol with minor modifications, such as the use of complex medium, cell densities in excess of 30 g/liter have been reported (Hsieh et al., 1988). Yields are highly strain- and protein-specific. Typical yields with this protocol are ~0.3 g/liter for intracellular products and ~100 mg/liter for extracellular products.

P. pastoris frequently yields intracellular products at 5% to 30% of cell protein. Cell density reaches 130 g dry weight/liter in the fermenter, of which >30% is protein; this is equivalent to at least 2 to 12 g/liter of product. As with *S. cerevisiae*, yields of secreted proteins are less predictable but frequently reach several hundred mg/liter at high cell density, and may constitute the only major protein in the culture medium. It is recommended that positive control strains—e.g., a strain expressing β-galactosidase—be used to ensure that the

BOULDER CENTRAL LIBRARY

induction conditions are optimal; control strains for *P. pastoris* can be obtained from Invitrogen.

Time Considerations

Small-scale *S. cerevisiae* inductions take a total of 2 days from inoculation of the starter culture to harvesting of the induced cells. Several inductions (e.g., 24) can be carried out at a time. *P. pastoris* inductions are generally carried out for 2 to 5 days; therefore, the procedure takes 3 to 6 days from start to finish, depending on the optimal length of induction for the particular protein.

S. cerevisiae fermentations typically take 4 days from start to finish: 2 days for inoculum development, 1 day for biomass accumulation on glucose, and ~1 day for galactose induction. *P. pastoris* 1-liter fermentations typically take 5 days from start to finish: 1 day to produce a starter inoculum, 2 days for biomass accumulation on glycerol, and up to 2 days for methanol induction. Some workers have reported that Mut^S inductions require a longer period, but this has not been the authors' experience. Half a day is required after the fermenter run to harvest and wash cells and to clean the fermenter.

Literature Cited

Aiba, S., Nagai, S., and Nishizawa, Y. 1976. Fed-batch culture of *Saccharomyces cerevisiae*: A perspective of computer control to enhance the productivity of Baker's yeast cultivation. *Biotechnol. Bioeng.* 18:1001-1016.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) 1995. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Bock Gu, M., Jung, K.-H, Park, M.N., Shin, K.S., and Kim, K.H. 1989. Production of HbsAg by growth-rate control with recombinant *Saccharomyces cerevisiae* in fed-batch. *Biotech. Lett.* 11:1-4.

Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. 1991a. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology* 9:455-460.

Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K., and Henwood, C.A. 1991b. Production of mouse epidermal growth factor in yeast: High-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* 105:205-212.

Cregg, J.M., Tschopp, J.F., Stillman, C., Siegel, R., Akong, M., Craig, W.S., Buckholz, R.G., Madden, K.R., Kellaris, P.A., Davies, G.R., Smiley, B.L., Cruze, J., Torregrossa, R., Velicelebi, G., and Thill, G.P. 1987. High-level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. *Bio/Technology* 5:479-485.

Cregg, J.M., Vedvick, T.S., and Raschke, W.C. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Bio/Technology* 11:905-910.

Fieschko, J.C., Egan, K.M., Ritch, T., Koski, R.A., Jones, M., and Bitter, G.A. 1987. Controlled expression and purification of human immune interferon from high-cell-density fermentations of *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 29:113-1121.

Hsieh, J.-H., Shih, K.-Y., Kung, H.-F., Shiang, M., Lee, L.-Y., Meng, M.-H., Chang, C.-C., Lin, H.-M.., Shih, S.-C., Lee, S.-Y., Chow, T.-Y., Feng, T.-Y., Kuo, T.-T., and Choo, K.-B. 1988. Controlled fed-batch fermentation of recombinant *Saccharomyces cerevisiae* to produce hepatitis B surface antigen. *Biotechnol. Bioeng.* 32:334-340.

Laroche, Y., Storme, V., De Meutter, J., Messens, J., and Lauwereys, M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technology* 12:1119-1124.

Pirt, S.J. 1985. Batch cultures with substrate feeds. *In* Principles of Microbe and Cell Cultivation, pp. 211-222. Blackwell Scientific, Oxford.

Romanos, M.A., Clare, J.J., Beesley, K.M., Rayment, F.B., Ballantine, S.P., Makoff, A.J., Dougan, G., Fairweather, N.F., and Charles, I.G. 1991. Recombinant *Bordetella pertussis* pertactin (P69) from the yeast *Pichia pastoris*: High-level production and immunological properties. *Vaccine* 9:901-906.

Romanos, M.A., Scorer, C.A., and Clare, J.J. 1992. Foreign gene expression in yeast: A review. *Yeast* 8:423-488.

Scorer, C.A., Clare, J.J., McCombie, W.R., Romanos, M.A., and Sreekrishna, K. 1994. Rapid selection using G418 of high-copy-number transformants of *Pichia pastoris* for high-level foreign gene expression. *Bio/Technology* 12:181-184.

Sonnlietner, B. and Kappeli, O. 1986. Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: Formation and verification of a hypothesis. *Biotechnol. Bioeng.* 28:927-937.

Sreekrishna, K., Nelles, L., Potenz, R., Cruze, J., Mazzaferro, P., Fish, W., Motohiro, F., Holden, K., Phelps, D., Wood, P., and Parker, K. 1989. High-level expression, purification, and characterization of recombinant human tumour necrosis factor synthesized in the methylotrophic yeast *Pichia pastoris*. *Biochemistry* 28:4117-4125.

Thill, G.P., Davis, G.R., Stillman, C., Holtz, G., Brierley, R., Engel, M., Buckholtz, R., Kinney, J., Provow, S., Vedvick, T., and Seigel, R.S. 1990. Positive and negative effects of multi-copy

EDU IAS CENTRAL UNIVERSITY LIBRARY

integrated expression vectors on protein expression in *Pichia pastoris*. *In* Proceedings of the 6th International Symposium on Genetics of Microorganisms, Vol. II (H. Heslot, J. Davies, J. Florent, L. Bobichon, G. Durand, and L. Penasse, eds.) pp. 477-498. Société Française de Microbiologie, Paris.

Wang, H.Y., Cooney, C.L., and Wang, D.I. 1979. Computer control of bakers' yeast production. *Biotechnol. Bioeng.* 21:975-995.

Key References

Cregg et al., 1993. See above.

An excellent review of P. pastoris expression technology.

McNeil, B. and Harvey, L.M. eds. 1990. Fermentation: A practical approach. Oxford University Press, Oxford.

A detailed laboratory manual covering all aspects fermentation, especially useful to those new to fermentation technology.

Romanos et al., 1992. See above.

A detailed review covering all aspects of yeast expression, including vector systems, factors affecting foreign gene expression, secretion, alternative yeasts (including P. pastoris), and fermenter scale-up.

Romanos, M.A., Scorer, C.A., and Clare, J.J. 1995. Expression of cloned genes in yeast. *In* DNA Cloning 2, Expression Systems: A Practical Approach (D.M. Glover and B.D. Hames, eds.) pp. 123-167. IRL Press, Oxford.

A detailed practical review of the basic strategies and techniques for high-level expression of genes in yeast including approaches for diagnosing and solving problems.

Contributed by Michael A. Romanos and Jeffrey J. Clare
Wellcome Research Laboratories
Beckenham, United Kingdom

Crawford Brown
British Bio-Technology Ltd.
Cowley, United Kingdom

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of Protein Expression by Mammalian Cells

In recent years, mammalian cells have been used in the production of recombinant proteins, antibodies, viruses, viral-subunit proteins, and gene-therapy vectors. In addition to being used in commercial biotechnology, mammalian cell systems have served as a means for examining fundamental aspects of gene replication, transcription, translation, and post-translational protein processing.

The availability of transformable cell lines, along with viral and plasmid-based mammalian-cell vector systems, has provided tools through which important aspects of mammalian gene function can be investigated. The following are typical uses for mammalian expression systems:

1. verification of a cloned gene product;
2. analysis of the effects of protein expression on cell physiology;
3. production and isolation of genes from cDNA libraries;
4. production of correctly folded and glycosylated proteins for assessment of biological activity in both in vitro and in vivo systems;
5. production of suitable quantities of proteins and glycoproteins for structural characterization of protein and carbohydrate moieties;
6. production of important clinically active viral surface antigens—e.g., prehepatitis B virus surface antigen (preS2 HBVsAg)—as well as therapeutic proteins—e.g., β-interferon, tissue plasminogen activator (tPA), erythropoietin (EPO), and Factor VIII; and
7. production of monoclonal antibodies.

Important features of mammalian cells include their ability to perform post-translational modifications and to secrete glycoproteins that are correctly folded and contain complex antennary oligosaccharides with terminal sialic acid. These covalent modifications may modulate the clinical efficacy of the protein (e.g., circulatory half-life and biospecificity) or result in properties that are of interest for biochemical characterization—e.g., with respect to structural stabilization, functional groups, and biological role. Mammalian-produced proteins are quality-controlled through a process whereby the progress of incompletely folded, misassembled, and unassembled proteins into the secretory pathway is selectively inhibited (Hurtley and Helenius, 1989). The correctly processed material progresses and is generally secreted as fully active protein. Another important feature of mammalian cell expression has been the ability to provide proteins bearing sialyl glycoforms closely resembling glycoproteins produced by humans. In contrast, *E. coli* does not possess any glycosylation machinery and yeast tends to produce hypermannose-type glycosylation. The yeast-produced glycoproteins are rapidly cleared from the circulation by the liver and also may induce adverse immunogenic responses in humans.

The post-translational modifications performed by mammalian cells include the proteolytic processing of the propeptide (Neurath, 1989) as well as glycosylation and carbohydrate trimming (Hirschberg and Snider, 1987). Other modifications may include γ carboxylation of glutamic acid residues (Suttie, 1985; Kaufman et al., 1986a); hydroxylation of aspartic acid and asparagine residues (Stenflo, 1989); sulfation of tyrosine residues (Baeuerle and Huttner, 1987); phosphorylation of proteins through cell receptor–protein interaction (Myers et al., 1994); fatty acid acylation (McIlhinney, 1990); and correct assembly of multimeric proteins. In contrast to what happens in mammalian cells, the production of intracellular recombinant proteins in *E. coli* leads to accumulation of the protein in reduced form, often as insoluble refractile bodies requiring subsequent denaturation and refolding.

This unit will briefly review the stages involved in protein production in mammalian cells using the stable-expression approach as outlined in Figure 5.9.1. Discussion will include choice of cell type, transfection of the host cells, methods for selection and amplification of transformants, and growth of cells at appropriate scale for protein production. Since post-transcriptional modification and intracellular protein transportation are important features of recombinant-protein production in mammalian cells, some description of these mechanisms is included.

CHOICE OF MAMMALIAN CELL HOST

Although a variety of mammalian cell hosts are available for protein production, only a small number have emerged as systems of choice for production of proteins to be used clinically. The most common of these cell hosts

Contributed by David Gray
Current Protocols in Protein Science (1997) 5.9.1-5.9.18
Copyright © 1997 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

are summarized in Table 5.9.1. The narrowing down of choices is largely due to the need for cell lines that: (1) are capable of continuous growth; (2) can be grown in suspension (in bioreactors); (3) have low risk of adventitious infection by potentially pathogenic viruses; (4) have genetic stability; and (5) can be readily characterized with respect to karyology, morphology, isoenzymes, and gene copy number. The existence of a variety of host-cell systems, the availability of viral or cDNA-based vectors, and the possibility of either stable or transient expression requires that the prospective user define an expression strategy based on ultimate goals. When the researcher's objectives require <1 mg of protein, transient expression in COS-7 cells is the relevant route. Transient expression in the COS-cell and vaccinia systems has been recently reviewed (Moss and Earl, 1991; Aruffo, 1997), and detailed protocols for construction of suitable vectors and protein expression by these systems can be found in those articles.

In transient expression a burst of production occurs in the host cell and is usually accompa-

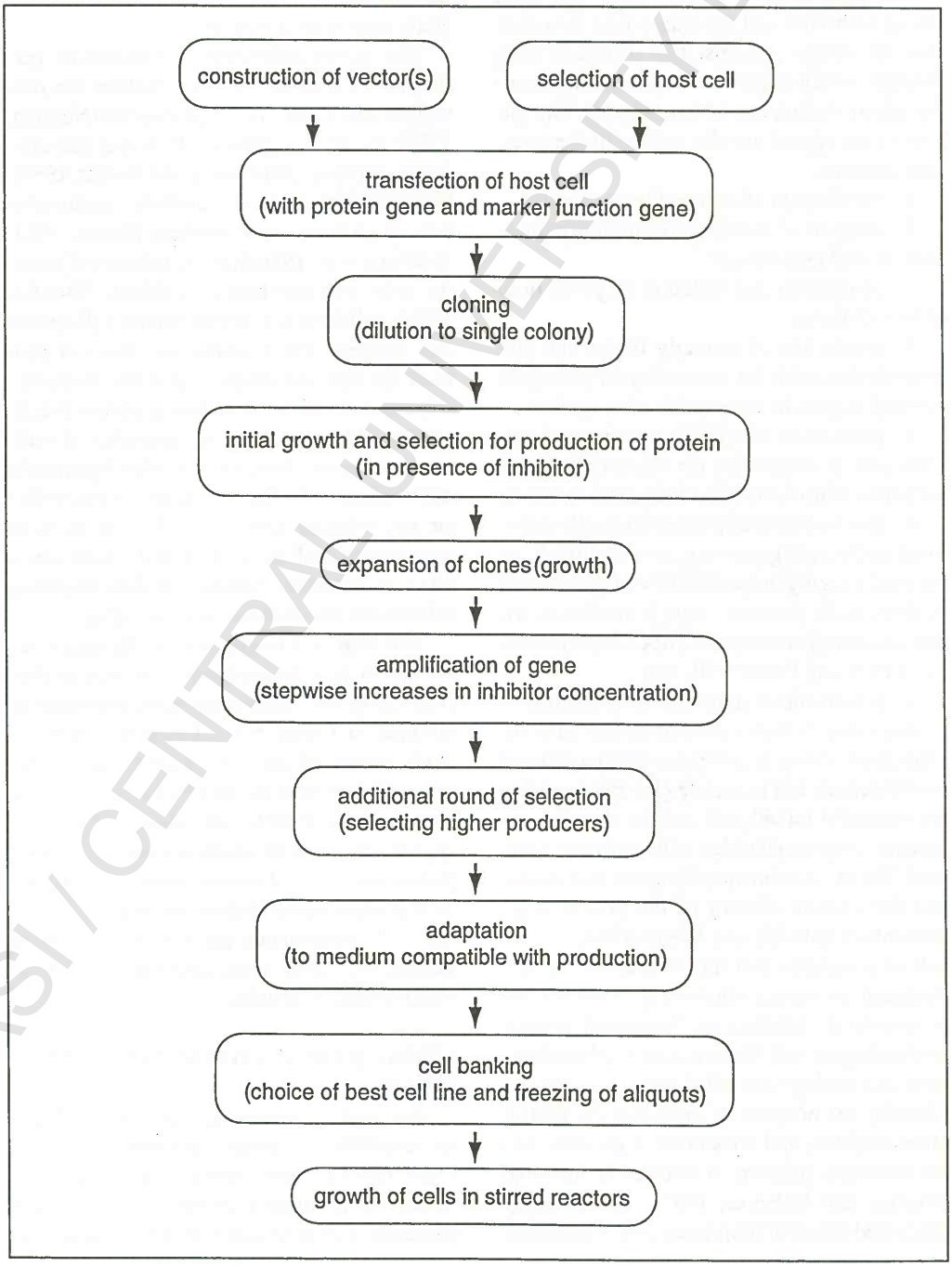

construction of vector(s)
selection of host cell
transfection of host cell
(with protein gene and marker function gene)
cloning
(dilution to single colony)
initial growth and selection for production of protein
(in presence of inhibitor)
expansion of clones (growth)
amplification of gene
(stepwise increases in inhibitor concentration)
additional round of selection
(selecting higher producers)
adaptation
(to medium compatible with production)
cell banking
(choice of best cell line and freezing of aliquots)
growth of cells in stirred reactors

Figure 5.9.1 Flow chart of activities in the development of a mammalian-cell expression system.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 5.9.1 Common Mammalian Cell Hosts

Cell line	Description	Growth	Utility	Source[a]
Human				
Namalwa	Burkitt's lymphoma–transformed lymphoblastoid cell	Large-scale suspension	Production of α interferon (e.g., by Burroughs Wellcome)	ATCC #CRL-1432
HeLa	Aneuploid cervical carcinoma cell	Small-scale suspension	Production of small quantities of research material (few mg)	ATCC #CCL-2
293	Transformed kidney cell	Small-scale suspension	Production of small quantities of research material (few mg)	ATCC #CCL-1573
WI-38	Human diploid normal embryonic lung cell	Attachment only	Host for virus production	ATCC #CCL-75
MRC-5	Human diploid normal embryonic lung cell	Attachment only	Hardier host for virus production—e.g., hepatitis A	ATCC #CCL-171
HepG2	Liver carcinoma transformed cell	Attachment	Small-scale evaluation of expression—e.g., HBVsAg	ATCC #HB-8065
Rodent				
3T3	Swiss mouse embryo fibroblast	Attachment	Used in testing transforming agents, expression evaluation	ATCC #CCL-92
L-929	Normal connective tissue fibroblasts	Attachment	Small-scale evaluations of expression	ATCC #CCL-1
Myeloma (e.g., NS/O)	Many types	Large-scale suspension	Monoclonal antibody production	ATCC (and commercial sources)
BHK-21	Baby hamster kidney cell	Large-scale suspension	Host for virus production or for stable gene integration	ATCC #CCL-10
CHO-K1	Chinese hamster ovary cell	Large-scale suspension	Used with glutamine synthetase system	ATCC #CCL-61
CHO DG44	Chinese hamster ovary cell	Large-scale suspension	Host for DHFR coamplification	L. Chasin[b]
CHO DXB11	Chinese hamster ovary cell	Large-scale suspension	Preferred host for DHFR coamplification	L. Chasin[b]
Monkey				
COS-7	Transformed African green monkey kidney cell	Small-scale attachment	Transient expression host	ATCC #CCL-1651
Vero	Normal African green monkey kidney cell	Large-scale attachment	Production of viruses	ATCC #CCL-81

[a]Abbreviation: ATCC, American Type Culture Collection (see *SUPPLIERS APPENDIX*).

[b]E-mail: *lac2@columbia.edu*

BCU IAS/CENTRAL UNIVERSITY LIBRARY

nied by death and rapid lysis of the cell. This presents the purification scientist with the challenge of fishing out the protein of interest, which may be present at ~5 μg/ml, from a soup of lysed cellular protein, nucleic acids, and viral particles. The yield of product during purification may be low as a result of the low titer and starting purity; however, when only small quantities of protein are required, transient expression in COS cells or the vaccinia system is a quick and suitable system to employ.

For production of larger quantities of protein, stable expression must be used because of the difficulty in scaling up transient expression into a bioreactor system. Stable expression in Chinese hamster ovary (CHO) cells has previously been described (Kingston et al., 1993).

Some cell lines have successfully been used as hosts in production of viruses and useful proteins. These include the human embryonic lung cell line MRC-5 and the normal embryonic cell line WI-38. These cells are attachment growth–dependent and have a finite lifespan of ~50 generations, after which time the cells enter a senescent phase and begin to die.

Cell lines that have undergone transformation by virus or that have experienced alteration of chromosomes can exhibit a capacity for infinite growth as well as an ability to grow in suspension. Examples of such cell lines are HeLa cells from a human cervical cancer and Namalwa cells from a human lymphoma. There has been some reluctance to use these cell lines in production of clinical agents because of the possibility of transfer of tumorigenic agents to the product. For production of larger amounts of recombinant protein, stable expression is required and has been primarily performed in CHO cells, baby hamster kidney (BHK-21) cells, or myeloma cells (e.g., NS/O). These cell lines are capable of indefinite growth on a large scale and are suitable hosts for stable integration of heterologous DNA.

Stable expression results from integration into the host-cell genome of the gene for the expression of the heterologous protein. The integrated gene is transcribed efficiently and the protein is expressed persistently over many generations by the host cell. A stable producer-cell line typically provides ~1 to 10 mg of secreted protein per 10^9 cells per day (specific productivity). For monoclonal antibody production, productivity levels of 35 to 100 mg per 10^9 viable cells per day in myeloma cells (Bebbington et al., 1992; Shitara et al., 1994) and 15 to 110 mg per 10^9 viable cells per day in CHO cells (Page and Sydenham, 1991) are obtainable. Such levels may allow secreted-antibody titers of 1 to 1.5 mg/liter to be achieved in optimized large-scale systems; hence these systems have been popular in biotechnology. Where protein is accumulated intracellularly, it may reach little more than 0.1% to 0.5% of total cell protein. Thus, the focus has tended to be on secretion of protein where significant amounts of product are obtainable.

The obvious drawback associated with stable expression in mammalian cells is the time required to obtain stable productive cell banks (months) versus that required to obtain large amounts of non-post-translationally modified proteins in *E. coli* or large quantities of high-mannose-containing glycoprotein using yeast and insect cells (2 to 3 weeks). The advantages of obtaining the more complex proteins from mammalian cells include the ability to understand how these molecules interact in biological models both in vitro and in vivo. Using mammalian systems, much knowledge has been gained in the understanding of post-translational modification, protein quality control, and protein translocation. This knowledge ultimately facilitates improved expression in mammalian hosts.

Stable expression requires the transfer of the foreign DNA, along with the relevant DNA-based signals for transcription by RNA polymerase II, into the chromosomal DNA of the host cell (e.g., CHO). A suitable plasmid vector is required to carry the functions through to the host DNA. Vectors are commercially available from a variety of vendors—e.g., Invitrogen, Clontech, and Promega (see *SUPPLIERS APPENDIX*). A large variety of vectors are described in Pouwels et al. (1988). Suitable replicons contain 5 to 10 kb of DNA and usually are designed to contain phenotypic markers for selection in *E. coli* (e.g., Ap^r) and CHO cells ($DHFR^+$). Other features of appropriate plasmids include a promoter-distal cloning site along with other cloning sites; a replication origin for *E. coli*; a polyadenylation sequence from SV40 or bovine growth hormone; a eukaryotic origin of replication (e.g., SV40 or oriP); the gene for expression of the protein in a site proximal to the promoter; and a promoter and associated elements. Commonly used constitutive promoters for CHO cell applications are the SV40 early promoter (Moreau et al., 1981; Neuhaus et al., 1984), the adenovirus late promoter (McKnight and Tjian, 1986), and the cytomegalovirus (CMV) early promoter (Boshart et al., 1985). Transcriptional enhancers such as the human CMV enhancer (Boshart et al.,

RCU IAS/CGPM UNIVERSITY LIBRARY

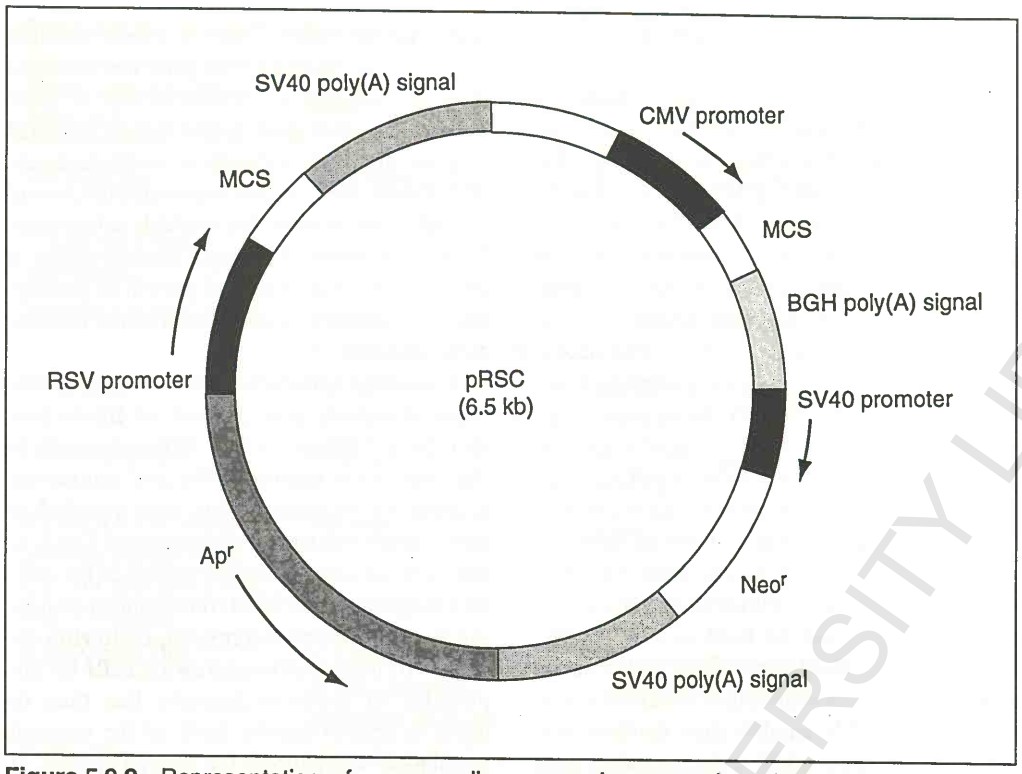


Figure 5.9.2 Representation of a mammalian expression vector based on the vector pRSC, published by Tsang et al. (1997); adapted with permission. The plasmid has two multiple cloning sites (MCSs) and three promoters. One promoter drives the neomycin-resistance protein while the remaining two MCSs have specific promoters and can be used to incorporate an amplification marker (DHFR) and a gene for the protein of interest.

1985) can enhance transcriptional activity by 10- to 100-fold and may be incorporated into the vector. These features are illustrated in Figure 5.9.2.

CHO cells have been used extensively as a host for stable expression of proteins. A number of CHO mutants have been developed that provide the user with tools to examine the synthesis of DNA, RNA, and protein as well as protein secretion, protein glycosylation, and intermediary metabolism. The most popular CHO sublines (also see Urlaub and Chasin, 1980) are CHO DXB11 (*dhfr*$^+$/*dhfr*$^-$) and CHO DG44 (*dhfr*$^-$/*dhfr*$^-$). The DXB11 cell was derived from CHO-K1 in 1978 by Lawrence Chasin. The CHO-K1 cells were originally derived in the 1950s, and the cells that were used in Chasin's laboratory for the development of the DXB11 line were obtained from Ted Puck and Fa-ten Kao at the Eleanor Roosevelt Cancer Institute in Denver in 1970. The DG44 cell line contains a double mutation in the *dhfr* genes and is not capable of natural reversion to the *dhfr*$^+$ phenotype. The DG44 cell line was derived from CHO pro3$^-$ cells by Chasin in 1982. The CHO pro3$^-$ was derived from the cell line established in the 1950s and is sometimes referred to as CHO Toronto, because it was used extensively in that city by Louis Siminivitch and colleagues. Both these cell lines can be obtained from Dr. Lawrence Chasin of Columbia University (e-mail: *lac2@columbia.edu*). Other mutants have been produced, and a brief list is included in Wirth and Hauser (1993).

TRANSFECTION, SELECTION, AND AMPLIFICATION

For introduction of genetic material into mammalian cells, three types of vectors are available—viruses, plasmid vectors with a replicon for animal cells, and plasmid vectors without a replicon for animal cells. Most viruses that replicate in mammalian cells ultimately cause death and lysis of the host cell, so expression of virus-borne genes is transient and can only be obtained over a short period prior to lysis. Retroviruses are capable of integrating genetic material into the chromosome of host cells—i.e., stable transformation. As with viruses, the use of plasmid vectors can result in either stable or transient expression of a gene product; the plasmid vectors without a mam-

BOSTON PUBLIC LIBRARY

malian cell replicon would require stable integration for expression of gene product.

Stable production of foreign proteins by mammalian cells requires use of a host cell line that is capable of incorporating the heterologous-gene DNA into a chromosome. In addition to the need for successful stable integration, a selection method is needed to fish out the stable transformants from the large population of cells that may contain unstable, intracellular heterologous gene copies. This necessitates cotransfection of a gene conferring some unique phenotypic property. Resistance to cytotoxic drugs is the most frequently applied method for selection of stable transformants. Normally, a gene for recessive drug resistance would be cotransfected into a host cell deficient in that particular activity. The most common approach is the use of dihydrofolate reductase (DHFR), which can be used as a selectable marker in DHFR-deficient CHO cells (Kaufman and Sharp, 1982). Other common transfectable selectable amplification markers that can be used with CHO cells are adenosine deaminase (in presence of inhibitors of ADA or cytotoxic levels of adenine or its analogs) in CHO DXB11 (DHFR$^-$; Yeung et al., 1983; Kaufman et al., 1986b); carbamyl phosphate synthetase–aspartate transcarbamylase–dihydroorotase in CHO UrdA mutant cells (de Saint Vincent et al., 1981); asparagine synthetase in CHO N3 (AS$^-$) cells (Cartier and Stanners, 1990); glutamine synthetase in any CHO cell (Bebbington and Hentschel, 1987); ornithine decarboxylase in CHO C55.7 (ODC$^-$; Chiang and McConlogue, 1988); multiple drug resistance in any CHO cell (Kane et al., 1989); and a mutant DHFR (Leu 22 changed to Arg 22) which can be used in any CHO cell host (Simonsen and Levinson, 1983). See Mortensen et al. (1997) for additional discussion of the use of selectable markers with mammalian cells.

It has been reported that a highly variable but typically large fraction (20% to 100%) of CHO clones established through calcium phosphate– or lipofection-mediated cotransfections (of *dhfr* and heterologous protein genes) coexpressed the dual-transfected expression vectors (Wurm and Petropoulos, 1994). It is likely that integration occurs at random sites. Given that the mammalian-cell genome is large (>10^9 bp), has 20% to 30% of the DNA as single-copy, and has only 0.1% of the genomic DNA containing coding sequences, it is critical to exert selective pressure to screen out the large number of nonproductive clones. Further amplification in the presence of selective pressure is necessary to obtain clones in which multiple copies of the heterologous gene are obtained through coamplification with the *dhfr* or other selectable marker gene. In this respect it is wise to screen for high producers at moderate levels of methotrexate (<1 μM) to prevent enrichment of high-copy-number but unstable integrands. Stability is assessed through determination of productivity after extended growth or passaging in the absence of selective pressure in complete medium.

Following transfection of cells using calcium phosphate precipitation or lipofection (Kaufman, 1990a), ~5% to 50% of the cells in the population acquire DNA and transiently express the encoded protein over a period of days before unstable, nonintegrated DNA is lost from the cells. A smaller portion of the cells will integrate the DNA chromosomally, resulting in stable protein expression. Following selection of these stable-expressing cells by application of selective pressure, less than or equal to approximately 0.1% of the original population will express the foreign protein. It is thus imperative to use selective pressure in the production of the early stable-expressing clones to prevent overgrowth with nonproducing cells. For CHO *dhfr*$^-$ host–based production, the cells are grown in medium deficient in nucleosides (e.g., α-MEM medium from Life Technologies), containing serum that has been dialyzed (to remove low-molecular-weight nucleosides such as hypoxanthine), and supplemented with 0.1 to 1.0 μM methotrexate. In deficient medium, the cell's requirement for purine and pyrimidine nucleosides is satisfied by the action of heterologous dihydrofolate reductase, which converts folate from the medium to tetrahydrofolate (FH_4). FH_4 and serine are then used by cellular serine hydroxymethyl transferase to provide methylene-FH_4 which can go on to donate C1 units in the reactions required for inosinic acid and thymidine production. Cells that acquire the capacity to express *DHFR* can be selected by medium deficient in nucleosides. Stepwise increases in the methotrexate concentration with increased passaging in deficient medium will select for cells containing high copy number of *DHFR* (Kaufman, 1990b).

PROTEIN TRANSLATION, QUALITY CONTROL, AND COVALENT MODIFICATION

Mammalian cells are complex factories with large numbers of compartments, each with specialized functions. Synthesis of proteins begins

Table 5.9.2 Protein Translocation and Post-Translational Modification Signals

Protein-based signal sequence	Related cell function	Example	Reference
$KDEL_{COOH}$	Retention in ER	BiP, protein disulfide isomerase, prolyl isomerase	Pelham and Bienz (1982)
Transmembrane sequences	Golgi retention	Coronavirus matrix protein	Mayer et al. (1988); Machamer and Rose (1987)
Mannose-6-phosphate	Added in Golgi for transport to lysosome	Lysosomal hydrolases	Reitman and Kornberg (1981)
SKL_{COOH}	Peroxisomes	Luciferase acyl coenzyme A oxidase	Keller et al. (1987); Miyazawa et al. (1989)
Leader peptide	Mitochondria	CoXIV (cytochrome *c* oxidase subunit IV)	Hurt et al. (1984)
RRNRRRRW (other specific signals exist, e.g., for p53 and SV40 late T antigen)	Nucleus	Rev	Malim et al. (1989)
Asn-X-Ser(Thr)	N-linked glycosylation in ER (dolichol precursor addition)	Many normal secreted and recombinant proteins	Hirschberg and Snider (1987)
O-Ser/Thr	Golgi-based O-glycosylation	Includes some recombinant proteins—e.g., IL-2, tPA	Hart et al. (1988); Conradt et al. (1990)
Cys (ER) Cys (Golgi)	Palmitoylation Myristoylation	Membrane anchoring—e.g., HBV preS1	McIlhinney (1990)
Tyr (Golgi)	Sulfation	Biological function	Rosenquist and Nicholas (1993); Han and Martinage (1992)
Tyr (cell-surface kinases)	Phosphorylation	Signal transduction and some viral proteins—e.g., HBV core	Myers et al. (1994); Albin and Robinson (1980); Aitken (1996)
Asn, Asp (Golgi)	Hydroxylation	Biological function	Stenflo et al. (1989)
Glu (Golgi)	γ-carboxyglutamate	Calcium binding—e.g., blood-cascade factors	Suttie (1985); Kaufman et al. (1986a)

in the nucleus with transcription and progresses to the cytoplasm, where translation and translocation into the endoplasmic reticulum (ER) occurs. The protein can contain sequences (motifs) that will order specific post-translational modifications as the protein progresses through the ER into the Golgi and before reaching the final destination either as a secreted protein or as a specialized resident cellular protein (see Table 5.9.2). In general, initiation of gene-transcription is the rate-limiting step. The rate of transcription elongation is constant until the movement of RNA polymerase II is stopped by termination signals. It follows that a highly expressed gene would be distinguished by the presence of a high density of polymerase molecules, resulting from high-frequency initiation.

For recombinant proteins that are destined for secretion, the translocation of expressed protein follows the so-called constitutive route (Walter and Lingappa, 1986), which is shown in part in Figure 5.9.3. A signal-recognition particle (SRP) binds to the signal sequence of the growing peptide when it protrudes into the cytoplasm, thereby halting further synthesis of the protein until the ribosome-SRP complex has bound to the docking protein (SRP receptor) at the ER membrane. Protein synthesis then continues and the peptide chain progresses into the lumen of the ER. At this point the amino-terminal signal sequence is cleaved off. Signal peptides encompass 13 to 30 amino acids. They usually have a net positive charge at the N terminus, a core of nine or more hydrophobic amino acids, and a cleavage site (Von Heijne, 1983, 1984). The tissue plasminogen activator

BOSTON CENTRAL UNIVERSITY LIBRARY

(tPA) leader peptide has been used for efficient translocation (Burke et al., 1986). The protein is folded and glycosylation is initiated in the ER. Detailed discussion of the intracellular factors involved in protein translocation are given in a number of excellent reviews (Gething and Sambrook, 1992; Rothman and Orci, 1992; Wirth and Hauser, 1993; Rothman, 1994).

The protein begins to fold cotranslationally in the ER. The ER provides an environment optimized for protein folding and is distinguished by a relatively high concentration of calcium (Baumann et al., 1991). The folding and stability of some proteins is calcium-dependent (Lodish and Kong, 1990). Another significant feature of the ER is that its redox potential is sufficiently high to enable disulfide-bond formation (Hwang et al., 1992). It appears

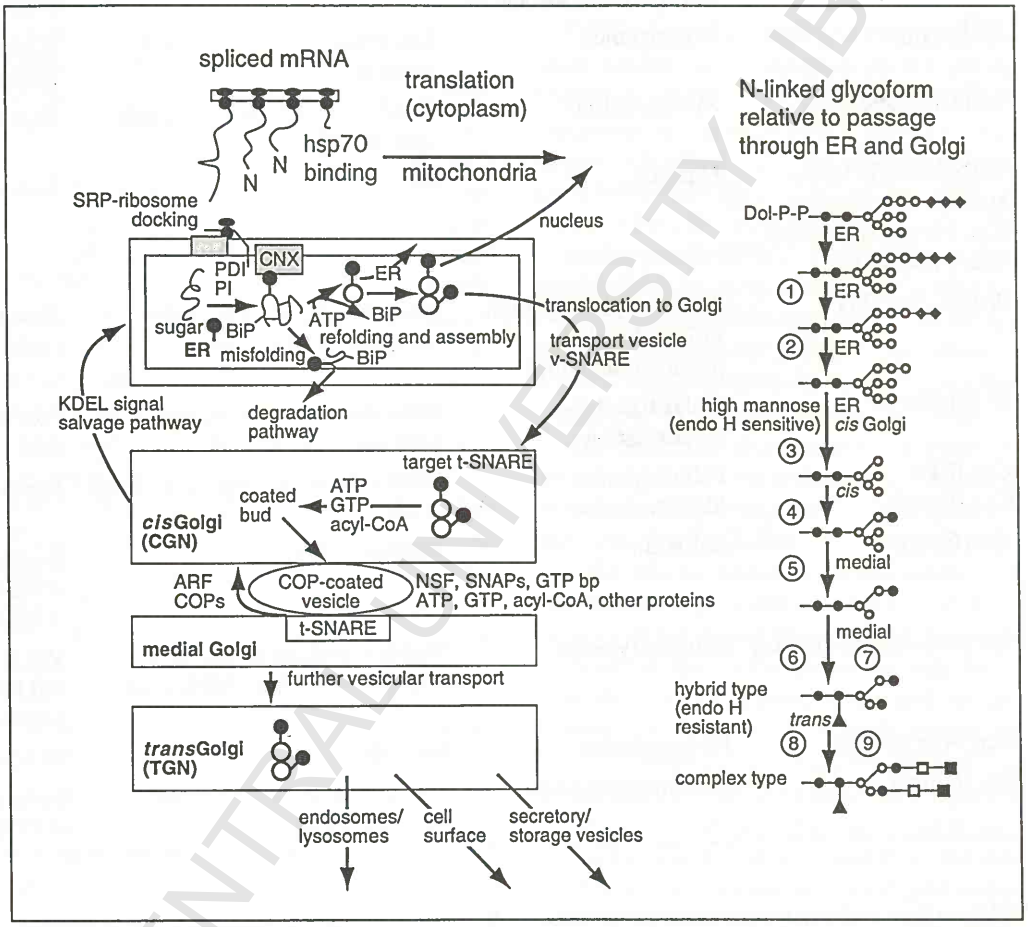


Figure 5.9.3 Intracellular protein translation and cisternal transport. Cytoplasmic translation results in a nascent polypeptide that is chaperoned by the hsp70 protein, which may direct the ribosome-peptide complex to the mitochondrion or the ER upon N-terminal peptide signaling. A signal-recognition particle (SRP) binds to the signal sequence, halting translation until it has docked at the SRP receptor. The hsp70 protein maintains the protein in a translocation-competent state, enabling membrane protrusion. Inside the ER the protein undergoes refolding assisted by protein disulfide isomerase (PDI), BiP, and calnexin (CXN). Initial N-glycosylation takes place cotranslationally in the ER. Misfolded protein, associated with BiP, is retained and eventually degraded. Transport of folded, assembled protein between the ER and Golgi cisternae occurs through a series consisting of budding, vesicle formation, and specific binding to target cisternae surfaces, through to the *trans* Golgi (TGN), where distribution to particular destinations occurs. **Abbreviations:** ARF, ADP-ribosylation factor; COPs, a family of coat proteins; Dol, dolichol; endo H, endoglycosidase H; GTP bp, GTP binding protein; NSF, NEM-sensitive cytosolic factor (identical to yeast sec 18p); SNAPs, three soluble NSF attachment proteins; v-SNARE and t-SNARE, SNAP receptors. **Symbols:** solid circles, N-acetyl glucosamine; open circles, mannose; solid diamonds, glucose; open squares, galactose; solid squares, sialic acid; solid triangles, fucose; circled 1, α-glucosidase I; circled 2, α-glucosidase II; circled 3, α-1,2-mannosidase; circled 4, GlcNAc transferase I; circled 5, α-mannosidase II; circled 6, GlcNAc transferase II; circled 7, α-1,6-fucosyl transferase; circled 8, β-1,4-galactose transferase; and circled 9, α-2,3-sialyl transferase.

BCU IAS CENTRAL UNIVERSITY LIBRARY

that glutathione is the redox buffer and that the redox potential of the ER is 20 to 100 times more oxidizing than that of the cytosol as indicated by the GSH/GSSG ratios (between 1:1 and 1:3 in the ER as compared to between 30:1 and 100:1 in the cytosol; Hwang et al., 1992). The ER contains soluble folding enzymes such as protein disulfide isomerase (PDI) and prolyl isomerase, which serve to aid the refolding process. Another group of proteins known as chaperones (Ellis and Hemmingsen, 1989) are present in the ER and appear to assist in protein folding and to ensure that incorrectly folded protein is not released from the ER, thus playing a quality-assurance role for newly synthesized protein in the cell. The best-studied chaperone is the glucose response protein BiP (or GRP 78), a member of the hsp70 family of proteins (Munro and Pelham, 1987; Hendershot et al., 1988). An ER integral membrane protein called calnexin (Bergeron et al., 1994) also appears to function in quality control of protein folding in the ER and affects the retention of incorrectly folded protein. Calnexin is unrelated to the heat-shock chaperone Hsp60, Hsp70, and Hsp90 families.

As the protein is folding in the ER, N-linked oligosaccharides are added to the nascent chain via dolichol phosphate. Intramolecular disulfide bonds are formed with the assistance of protein disulfide isomerase (PDI), and the molecular chaperone BiP/GRP78 associates with the refolding peptide chain. Multiple cycles of BiP binding and release occur in a process that requires ATP. Protein folding occurs with high efficiency and with a half-time on the order of 1 min. Pulse-chase studies of hepatitis B surface antigen (HBVsAg) secretion in mouse fibroblast cells (Huovila et al., 1992) showed that the formation of HBVsAg dimers in the ER occurred rapidly, within 2 min, while formation of higher-order oligomers occurred with a half-time of 2 hr. The study also provided evidence that the rate-limiting step was likely to be passage from the ER into the Golgi, as mature intracellular glycosylated HBV was not detected. The higher-order assembly of subunits or assembly of multimeric proteins begins in the ER and is likely completed in the so-called 15° compartment (named because of accumulation of molecules there at low temperature). This appeared to be the case for HBVsAg (Huovila et al., 1992). When a monomeric subunit is extensively folded, the BiP dissociates, and for multimeric proteins such as viral membrane proteins, subunit-subunit recognition and assembly occurs prior to transport into the Golgi complex (Doms et al., 1993). Only correctly folded protein progresses from the ER to the Golgi apparatus. Protein may be misfolded as a result of inefficient folding, mutations, heat shock, energy depletion, or addition of exogenous reductants such as dithiothreitol. These misfolded proteins usually form aggregates and typically display permanent and stable association with BiP. Aggregates are eventually degraded unless they are salvaged through a pathway that facilitates refolding from the misfolded protein species under specific conditions—e.g., altered temperature for temperature-sensitive folding mutants or where the ER environment is experimentally altered.

Proteins that have successfully folded in the ER are destined for secretion to the plasma membrane or the intracellular compartments, including the ER itself. Proteins destined to be retained in the ER possess the C-terminal signal sequence KDEL (Munro and Pelham, 1987). Proteins containing the KDEL signal that escape the ER are efficiently salvaged from the Golgi through a divalent cation–dependent receptor-based salvage pathway (Rothman and Orci, 1992). Proteins destined to be integrated into the ER membrane have a signal consisting of two lysine residues separated by three and four or five amino acids apart from the C terminus (Jackson et al., 1990). Signals for Golgi retention appear to involve histidine and cysteine within the membrane-spanning domain (Aoki et al., 1992).

N-linked oligosaccharide assembly begins in the ER compartment (Kornfeld and Kornfeld, 1985). In brief, N-linked glycosylation begins with synthesis of a lipid-linked oligosaccharide moiety—$Glc_3Man_9GlcNAc_2$-P-P-dolichol—which is transferred to the nascent peptide chain at the relevant Asn-X-Ser/Thr sequon site (i.e., peptide sequence recognition site). Truncations of the initial oligosaccharide structure follow this initial step in the ER. The folded and primary-glycosylated protein is transported to the Golgi, where maturation of the carbohydrate chains takes place. O-glycosylation is thought to occur after the N-glycosylation in the ER or the early part of the Golgi.

Differences in glycosylation of a given protein are a result of divergence in the Golgi-based processing of the oligosaccharides in different cell types (Jenkins et al., 1996). The human lymphoblastoid Namalwa cell line performs O- and N-linked glycosylation of recombinant tPA efficiently to produce human-type glycosylation characteristics (Okamoto et al., 1991; Khan et al., 1995). Most CHO cell lines

B. C. U. „M. I. ... 'AŞI

used in production of recombinant proteins do not possess active α1,3-galactosyltransferase, although the gene is present in these cells, which make low levels of galactosamine neuraminic acid. CHO and BHK cells also lack a functional α2,6-sialyltransferase enzyme and synthesize exclusively α2,3–linked terminal sialic acids via α2,3-sialyltransferase, in contrast to human cell lines, which contain both enzymes (Lee et al., 1989). In CHO cells the dominant sialic acid found is N-acetyl neuraminic acid.

Transport of the protein through the secretory and endocytic pathways is mediated by small vesicles that bud off from one compartment and fuse to the next compartment (Rothman and Orci, 1992). The *cis* Golgi stack (CGN) and *trans* Golgi stack (TGN) are defined areas of the entire Golgi network. The CGN and TGN constitute the entry and exit faces, respectively, and are involved in sorting and distribution of the protein. The CGN includes the KDEL salvage pathway for ER-based proteins. The TGN is where proteins with differing final destinations—e.g., lysosomes, secretory vesicles, and plasma membrane—diverge. The Golgi consists of stacks of cisternae where carbohydrate-chain trimming and elongation as well as other covalent modifications occur. The mechanism of flow through the Golgi cisternae appears to involve assembly of coated vesicles, which then transfer to the next compartment with concomitant loss of the coat. The loss of coat is linked to GTP-dependent hydrolysis, while subsequent fusion at the destination cisternae is ATP-dependent and also involves an N-ethylmaleimide (NEM)-sensitive enzyme, Sec18p. The specificity of vesicle targeting involves a docking mechanism based on binding of soluble NEM-sensitive attachment protein (SNAP) to a receptor designated as SNARE. Heterogeneity in the SNARE family allows for specialization and the ability to direct a given protein in the appropriate direction by selective fusion to progressive cisternae and ultimately the destination membrane.

Although it is usually desirable to obtain secretion of protein out of the cell prior to purification, in some cases the goal may be to express a protein receptor that would be transported to the cell-surface membrane. This was described for the production of the erythropoietin EPO receptor (EPO-R) in transfected hematopoietic cells (Hilton et al., 1995). The polypeptide was mutated to improve EPO-R refolding in the ER and increase cell-surface expression. The protein was obtained as a crude extract from the cell culture medium following the completion of the cell culture process. The glycoprotein may then be purified for further physicochemical analysis or testing in biological assays and animal studies.

Some proteins are not normally secreted by mammalian cells. This is the case for ER-resident proteins such as BiP, prolyl isomerase, and protein disulfide isomerase, which contain the C-terminal ER-retention peptide sequence KDEL. Also, a number of proteins including receptors and signal-transduction proteins are expressed and retained in the cell or on the cell surface. In cases where specific inhibitors of glycosylation enzymes are added to the culture medium, proteins that are normally secreted may be poorly secreted and thus accumulate in the ER or Golgi of the cell (Hickman and Kornfeld, 1978). Release of these intracellular or surface proteins requires disruption of the cell using a method capable of breaking the cell membrane (homogenization); in cases where membrane association is expected, nonionic detergents at 0.1% to 0.5% (v/v) are used in the lysis buffer.

Covalent modifications of amino acids by mammalian cells are numerous and to some extent may be a feature of a particular cell type. For example the γ carboxylation of factor VIII, which is essential for clotting-cascade activity, can be accomplished in the BHK and human 293 cell lines. This microsome-based, vitamin K–dependent modification was nonfunctional in CHO cells that had been transformed with complementing genes for carboxylation (Wirth and Hauser, 1993). Other processing events occur during passage of the protein from the ER to the Golgi. These involve removal of certain residues from the carbohydrate chains—e.g., glucose and mannose trimming. Other sugars are added to the carbohydrate chains as the protein passes through the medial and *trans* Golgi—e.g., addition of *N*-acetylglucosamine, galactose-*N*-acetate, fucose, galactose, and sialic acid. O-linked glycosylation is important for stabilization of some proteins—e.g., low-density-lipoprotein (LDL) receptor (Kingsley et al., 1986).

GROWTH OF MAMMALIAN CELLS FOR PROTEIN EXPRESSION

Mammalian cell culture media are complex in comparison to the simple defined media used for propagation of bacteria and yeast. This complexity stems from the need to satisfy the many functions required for normal growth and

RCU IAS LIBRARY UNIVERSITY LIBRARY

metabolism of mammalian cells (Bettger and McKeehan, 1986). Nutrients must satisfy the requirements for growth, maintenance of the cell, and expression of products. Additional specific growth factors such as insulin or insulin-like growth factor (IGF) are usually required. Minimal nutritional requirements were first defined in the 1950s (Eagle, 1955). These early studies showed that thirteen amino acids (Arg, Cys, Gln, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Tyr, and Val), eight vitamins (biotin, folic acid, niacinamide, pantothenic acid, pyridoxine, riboflavin, and thiamine), and various salts and minerals (Na, K, Ca, Mg, Cl, PO_4, Fe, Zn, Se, and possibly Cu, Mn, Mo, and V, which are usually present as trace contaminants) are required. The commonly used CHO derivatives DXB11 and DG44 have an absolute requirement for proline that results from a genetic deficiency in the conversion of glutamic acid to glutamic γ-semialdehyde. Inorganic ions may function as catalysts or exert physiological effects. Glucose (5 to 20 mM) and glutamine (0.5 to 5 mM) are commonly used to provide the energy source for the cells. These are consumed at a rate proportional to the viable-cell concentration. Cultured mammalian cells do not use the complete tricarboxcylic acid (TCA) cycle, and the glucose-consumption rate is lower than the glutamine consumption rate. Consumption of glucose yields lactate and consumption of glutamine yields ammonia as the metabolic byproduct. Other nutrients are consumed at a rate proportional to the rate of increase in cell density. Yield-based uptake rates for amino acids by CHO DG 44 cells range from 0.1 mM tryptophan/10^9 cells to 0.8 mM arginine/10^9 cells at a specific growth rate (the rate of growth divided by the cell density) of 0.25 to 0.30 $days^{-1}$ (D. Gray, unpub. observ.), reflecting the relative incorporation frequency of the individual amino acids into general protein. CHO cells, like most mammalian cells, produce alanine as a result of glutaminolysis. Accumulation of ammonia, as a result of glutamine degradation in the medium and consumption by cells, may become inhibitory to cell growth and ultimately limit the density of the culture (Butler and Spier, 1984).

In addition, physiological parameters such as pH, osmolality, and redox potential must be kept within acceptable limits. The use of serum is often advocated and has been regarded as a source of important components such as hormones, lipids, and growth factors. These serum-borne ingredients provide some consumable nutritional components as well as mitogens or growth factors that overall are not consumed by the cells (Klinman and McKearn, 1981).

A further level of complication in mammalian-cell medium design occurs through the need to maintain balanced relationships between nutrients while maintaining the osmolality of the medium within the physiological constraints (290 to 380 mOsmol/kg).

Recent medium development has focused on reducing the need for high levels of serum (Jayme, 1991). Fetal bovine serum (FBS) is the most frequently used version of serum and can be obtained from HyClone (see *SUPPLIERS APPENDIX*). FBS costs ~$500 to $600 per liter and is a factor contributing to the high cost of producing commercial clinical lots of protein. There may be batch-to-batch variation in serum which necessitates screening by small T-flask or shake-flask growth studies with the chosen mammalian cell line. Serum contains ~35 mg/ml of protein, of which ~60% is albumin. Thus if 10% serum were added to medium, the level of initial exogenous protein in the media would be 3 to 5 mg/ml.

Growth of cells in low serum (<0.5% v/v) has been accomplished through step-wise adaptation of cells to growth in medium with low or zero serum. The important functions of serum are replaced by the addition of the lipid precursors choline, inositol, and ethanolamine, the polyamine putrescine, and recombinant insulin (Nucellin-Zn from Eli Lilly) at 1 to 5 mg/liter (Jenkins, 1990) along with $FeCl_3$/citrate (Keenan and Clynes, 1996). A number of serum-free media have been developed and are now commercially available from, e.g., BioWhittaker. Another approach to implementing protein-free media has been the genetic modification of CHO cells through incorporation of the heterologous genes coding for transferrin and IGF-1 (Pak et al., 1996). As a result of cellular expression of the IGF-1 and transferrin, the cell is then able to grow, without addition of exogenous proteins, in a modified DMEM/F12 medium supplemented with sodium selenite.

A major concern in producing recombinant protein in protein-free medium is the potential for proteolysis and glycosidase activity. Such activity results in processing of the polypeptide chain and degradation of the carbohydrate moiety (Gramer and Goochee, 1993; Teige et al., 1994). Serum contains endogenous protease inhibitors—e.g., α_2-macroglubulin (Barrett, 1981)—and for this reason may be a valuable reagent in its own right in cell culture media

UCSF LIBRARY

when the produced protein is susceptible to proteolysis. Sialidases retain significant activity at pH 7.0 and are active in cell culture supernatants. Caution should be taken to minimize cell lysis during the culture process if the goal is to minimize desialylation of the expressed protein. This may be achieved by harvesting the supernatants during the exponential-growth phase prior to the beginning of the decreasing-viability phase, chilling it, and adding sialidase inhibitors such as 2,3-dehydro-2-deoxy-*N*-acetylneuramic acid to the supernatant (Gramer et al., 1995). In batch culture in serum-free medium, cell viability may decrease rapidly at the end of culture, thereby releasing intracellular proteases and glycosidases into the culture medium, which may in turn cause rapid proteolysis and carbohydrate degradation of the product.

Production of sialylated forms of glycoproteins is important in clinical applications to avoid rapid clearance by asialo- and asialogalactoprotein receptors in the liver. The genes for some of the key enzymes in the glycosylation pathway have been transfected into CHO cells (Lee et al., 1989) in attempts to tailor the glycoforms. A number of CHO-glycosylation mutants have been made available that exhibit a defect in one or more steps in the pathway (Stanley, 1989). Erythropoietin (EPO) expressed in the CHO mutant Lec 3.2.8.1. exhibited less heterogeneity than EPO produced in wild-type CHO cells (Stanley, 1991).

Studies with CHO-320 cells in glucose-limited chemostat culture (using RPMI-1640 medium) at a low dilution rate of 0.008 day^{-1} (i.e., volume of fresh medium added to the reactor per day divided by the reactor medium volume) showed a deviation of the specific growth rate from the dilution rate and an increase in specific cell-death rate (i.e., rate of cell death per liter per day divided by the number of viable cells per liter), suggesting that a minimum specific growth rate of 0.011 day^{-1} existed (Hayter et al., 1993). In the same study, the specific rate of γ interferon (γ-IFN) production increased with the specific growth rate (μ), indicating growth-related product production. The pattern of γ-IFN glycosylation was similar at all growth rates except the lowest, where an increase in nonglycosylated γ-IFN was observed. This was further addressed using batch culture (Leelavatcharamas et al., 1994), where it was concluded that the growth-associated production was a reflection of a cell cycle–related phenomenon whereby lower growth rate resulted in a prolongation of the cell-cycle time with an increase in the relative frequency of the G_1 cells in the culture population. It was suggested that the G_1 phase is the least productive phase in the cell cycle of CHO cells. Similarly S phase–specific synthesis of dihydrofolate reductase in nontransfected CHO-K1 cells was observed in which the maximum peak of DHFR activity coincided with the maximum rate of DNA synthesis in late S phase (Mariani et al., 1981). The expression of recombinant β-galactosidase by the CMV late promoter in CHO βG72-16 was shown to be S-phase related (Gu et al., 1993). It is likely that CHO cells will generally exhibit an S-phase dependency for expression of foreign genes using the typical promoter elements such as the CMV late promoter, and that a direct correlation of specific heterologous protein production rate with specific growth rate will be observed (Bock et al., 1993; Gu et al., 1996). As growth rate increases, the ratio of $S/(G_1 + G_2/M)$ increases (Hooker et al., 1995) and the frequency of cells in productive S phase in the culture increases.

A given glycoprotein produced in cell culture by mammalian cells will present a population of molecules exhibiting differences in glycosylation. The existence of families of structurally related yet distinct carbohydrate chains makes structural analysis of a distinct species complicated, since it is difficult to separate and purify the structural variants. Variation may result from incomplete occupancy of a directed site on the protein and from carbohydrate-structure variations imparted to the cells via cell culture–based conditions. Furthermore the action of glycosidases in the extracellular environment, particularly during batch culture, may degrade the carbohydrate structure of secreted glycoprotein, leading to a wider range of carbohydrate structures in the population. Changes in the oligosaccharides of the glycoprotein may lead to diminished biological efficacy in vivo—e.g., through increased clearance—or may enhance unwanted antigenicity of the molecule in a clinical application (Goochee et al., 1991).

Antibodies produced in CHO cells in serum-containing medium exhibited more galactosylation compared to antibodies produced by cells grown in serum-free medium (Lifely et al., 1995). The culture-medium glucose concentration affected the degree of glycosylation of γ-IFN produced in CHO cells in continuous culture (Hayter et al., 1992). Lipid supplements, with or without lipid carriers such as lipoprotein, appeared to increase N-glycosylation-site occupancy in γ-IFN (Jenkins et al.,

RCU IAS LIBRARY

Table 5.9.3 Estimated Quantity of Foreign Protein Produced by a Given Production System[a]

System	Vol. of medium in system	Total cells in system (vol. × conc.)	Total product in system[b]
Batch T-flask (T-175):			
175 cm^2 attached growth in DMEM/F12/10% serum[c]	75 ml	1.8×10^7	0.5-1 mg
Batch shake flask (250 ml), 0.5-1% serum[c]	100 ml	2×10^8	2-4 mg
Batch 5-liter spinner flask[c]	1000 ml	1×10^9	10-20 mg
Batch 5-liter spinner flask fitted with O_2 frit sparger[c]	1000 ml	2×10^9	20-40 mg
Batch 15-liter bioreactor with pH and O_2 control[c]	10 liters	2×10^{10}	200-400 mg
Continuous-mode 15-liter perfusion reactor with pH and O_2 control and cell retention	5-20 liter perfusion per day	1×10^{11}	100-400 mg per day

[a]Quantities assume a specific productivity of 1 to 4 mg/10^9 viable cells per day.

[b]Production of recombinant intracellular protein may achieve ~0.1% of total cell protein.

[c]Batch culture would be performed for 4 days after inoculation of 1×10^5 viable cells/ml.

1994). In perfusion culture of BHK-21 cells producing a recombinant interleukin 2 (IL-2) mutant, nutrient limitations (i.e., glucose, amino acids, and dissolved O_2) led to short-term changes in the general level of glycosylation, but the specific type of sugar incorporated was largely unchanged (Gawlitzek et al., 1995).

Direct intervention in cell physiological processes has an impact on glycosylation. A decreased protein-synthesis rate following addition of cyclohexamide improved the glycosylation-site occupancy of recombinant prolactin produced by C127 cells (Shelikoff et al., 1994). Low concentrations of dithiothreitol, which prevent cotranslational disulfide-bond formation in the ER, led to complete glycosylation of tPA at a tripeptide sequon that normally exhibited variable occupancy (Allen et al., 1995). This suggests that refolding may in itself limit the accessibility of glycosylation sites.

Sodium butyrate is sometimes used to improve protein synthesis, but it can change glycosylation by inducing the GlcNAc-transferase involved in O-glycosylation (Datti and Dennis, 1993) and by increasing sialyltransferase activity in recombinant CHO cells (Chotigeat et al., 1994; Gebert and Gray, 1995).

SCALE OF OPERATION

Batch growth in shake flasks or spinner flasks may provide titers in stable producer cell lines of 10 to 50 mg/liter, thus yielding up to 30 to 150 mg of crude secreted product which may then be purified from the culture supernatant. Following purification, the researcher may recover ~10% to 30% of the protein. Thus, small laboratory-scale cell culture may provide 3 to 50 mg of pure protein (>90% purity). Larger quantities of material require use of bioreactors in which pH, dissolved O_2, temperature, and stirring may be controlled. Should the user wish to control the cell growth rate during the production of the product, a continuous reactor may be used at a fixed dilution rate. In cases where high cell density is desired, the mode of operation would be continuous medium perfusion with cell retention in the bioreactor. Tables 5.9.3 and 5.9.4 provide some guidance for the likely scale requirements for a given objective.

The bioreactor environment may influence the product quality. Mild hypoxia may decrease the specific activity of sialyl transferase and hence the sialylation of the oligosaccharide chain, as was observed with recombinant human follicle-stimulating hormone produced in CHO cells at 10% dissolved oxygen saturation (Chotigeat et al., 1994). It appears that pH values within the range 6.9 to 8.2 have no dramatic effect on glycosylation, but clearly optimal growth would require control of pH within a narrower range, typically 7.0 to 7.4.

Metabolic byproducts may be the limiting factors in achieving high cell density in perfu-

BOSTON UNIVERSITY LIBRARY

Table 5.9.4 Typical Protein Requirements for Various Analyses

Assay	Quantity of protein required	Comments
ELISA	<5 µg	Identification of activity possible with crude extract
Amino acid analysis and sequencing (including development of method)	20 µg	Pure protein required; thus starting requirement would be >100 µg in order to account for purification losses
Carbohydrate analysis	20 µg	Pure protein required
Extinction-coefficient determination	500 µg	Pure protein required; thus starting requirement would be 3-5 mg
Biological testing in animal-based model	50 mg	Must obtain suitable purity of product to prevent collateral effects on physiology due to contaminants
Human clinical trials	1-100 g	>95% purity; requires extensive testing to ensure safety

sion culture. Increases in ammonium ion concentration to 10 mM in the culture medium resulted in substantial growth inhibition in hybridoma cells (McQueen and Bailey, 1990) and appeared to inhibit growth in most cell lines. The likely cause of this is the transport of ammonium ion by membrane-bound ion pumps and ultimate acidification of the cytoplasm (Martinelle et al., 1996). Lower levels of ammonia (~2 mM) may exert influences on glycosylation (Andersen et al., 1994). One approach to eliminating this potential inhibition of cell growth is through adaptation to a nonammoniagenic medium. This medium would contain glutamate or 2-oxoglutarate in place of glutamine (Hassel and Butler, 1990). Lactate generated from glucose can also reach growth inhibitory levels at >15 mM. Optimization of medium and perfusion rate may enable reduction in the levels of lactate and ammonia, allowing higher cell density. Lowering the glucose level may reduce the yield of lactate, but concomitant limitation of glycosylation may occur. CO_2 may accumulate in high-density cultures, thereby exerting an inhibitory effect on the culture (Gray et al., 1996). Limitations that are due to CO_2 accumulation may be overcome by choosing an optimal sparging strategy.

Finally mammalian cells are weaker and larger than bacterial cells and limits to the degree of agitation are important to prevent cell damage and death. Where agitation is used in reactors or in sample handling, cell damage at the gas-liquid interface as well as bubble damage can be reduced by adding low levels—<0.1% (w/v)—of the surfactant pluronic polyol (Murhammer and Goochee, 1990).

SUMMARY

Both yeast and *E. coli* heterologous protein expression are very well defined. The components of these expression systems are easily obtained through many commercial vendors. Use of these systems enables the researcher to obtain large amounts of pure material in a short period of time. In contrast, mammalian expression systems are not comprehensively defined or understood and require a significant investment of time before heterologous protein is obtained. At the present time the mammalian expression systems are not available as turnkey, catalog-based systems. The user must define the components for a particular application. Although this complexity exists with the mammalian cell–based systems, they are uniquely important for the production of glycoproteins that possess oligosaccharides close to or identical to those of human origin. Proteins secreted by mammalian cells possess correct conformations and exhibit full biological activity. The ability to transfer foreign DNA into these cells enables a greater understanding of the mammalian cell transcriptional, translational, and post-translational machinery. The ability to incorporate foreign gene sequences into a mammalian-cell chromosome and the ability to grow mammalian cells in suspension in large-scale bioreactors has enabled the production of large quantities of pure clinically important proteins.

RCU IASI - CENTRUL UNIVERSITAR IBMPP

LITERATURE CITED

Aitken, A. 1996. Protein chemistry methods, post-translational modification, consensus sequences. *In* Protein LabFax (N.C. Price, ed.) pp. 253-285. BIOS Scientific Publishers, Oxford.

Albin, C. and Robinson, W.S. 1980. Protein kinase activity in hepatitis B virus. *J. Virol.* 34:297-302.

Allen, S., Naim, H.Y., and Bullied, N.J. 1995. Intracellular folding of tissue-type plasminogen-activator-effects of disulfide bond formation on N-linked glycosylation and secretion. *J. Biol. Chem.* 270:4797-4804.

Andersen, D.C., Goochee, C.F., Cooper, G., and Weitzhandler, M. 1994. Monosaccharide and oligosaccharide analysis of isoelectric focusing–separated and blotted g-CSF glycoforms using high pH anion exchange chromatography with pulsed amperometric detection. *Glycobiology* 4:459-467.

Aoki, D., Lee, N., Yamaguchi, N., Dubois, C., and Fukuda, M.N. 1992. Golgi retention of a trans-Golgi membrane protein, galactosyl-transferase, requires cysteine and histidine residues within the membrane-anchoring domain. *Proc. Natl. Acad. Sci. U.S.A.* 89:4319-4323.

Aruffo, A. 1997. Transient expression of proteins using COS cells. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.13.1-16.13.7. John Wiley & Sons, New York.

Baeuerle, P.A. and Huttner, W.B. 1987. Tyrosine sulfation is a trans-Golgi-specific protein modification. *J. Cell. Biol.* 105:2655-2664.

Barrett, A.J. 1981. α_2-macroglobulin. *Methods Enzymol.* 80:737-754.

Baumann, O., Walz, B., Somlyo, A.V., and Somlyo, A.P. 1991. Electron probe microanalysis of calcium release and magnesium uptake by endoplasmic reticulum in bee photoreceptors. *Proc. Natl. Acad. Sci. U.S.A.* 88:741-744.

Bebbington, C.R. and Hentschel, G.C.G. 1987. The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells. *In* DNA Cloning, Vol. 3 (D.M. Glover, ed.) pp. 163-168. IRL Press, Oxford.

Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D., and Yarranton, G.T. 1992. High level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Bio/Technology* 10:169-175.

Bergeron, J.J.M., Brenner, M.B., Thomas, D.Y., and Williams, D.B. 1994. Calnexin: A membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem. Sci.* 19:124-128.

Bettger, W.J. and McKeehan, W.L. 1986. Mechanisms of cellular nutrition. *Physiol. Rev.* 66:1-35.

Bock, G., Todd, P., and Kompala, D.S. 1993. Foreign gene expression (β-galactosidase) during the cell cycle phases in recombinant CHO cells. *Biotechnol. Bioeng.* 42:1113-1123.

Boshart, M., Weber, F., Gerhard, J., Dorsch-Hasler, K., Fleckenstein, B., and Schaffner, W. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* 41:521-530.

Burke, R.L., Pachl, C., Quiroga, M., Rosenberg, S., Haigwood, N., Nordfang, O., and Ezban, M. 1986. The functional domains of coagulation factor VIII:C. *J. Biol. Chem.* 261:12574-12578.

Butler, M. and Spier, R.E., 1984. The effects of glutamine utilization and ammonia production on the growth of BHK cells in microcarrier cultures. *J. Biotechnol.* 1:187-196.

Cartier, M. and Stanners, C.P. 1990. Stable high-level expression of a carcinoembryonic antigen–encoding cDNA after transfection and amplification with a dominant and selectable asparagine synthetase marker. *Gene* 95:223-230.

Chiang, T. and McConlogue, L. 1988. Amplification and expression of heterologous ornithine carboxylase in Chinese hamster ovary cells. *Mol. Cell. Biol.* 8:764-769.

Chotigeat, W., Watanapokasin, Y., Mahler, S., and Gray, P.P. 1994. Role of environmental conditions on the expression levels, glycoform pattern and levels of sialyltransferase for hFSH produced by recombinant CHO cells. *Cytotechnology* 15:217-221.

Conradt, H.S., Hofer, B., and Hauser, H. 1990. Expression of human glycoproteins in recombinant mammalian cells: Towards genetic engineering of N- and O-glycoproteins. *Trends Glycos. Glycotech.* 2:168-181.

Datti, A. and Dennis, J.W. 1993. Regulation of UDP-GlcNAc:Galβ 1-3 GalNacR β 1-6-N-acetylglucosaminyltransferase (GlcNAC to GalNAc) in Chinese hamster ovary cells. *J. Biol. Chem.* 268:5409-5416.

de Saint Vincent, B.R., Delbruck, S., Eckhart, W., Meinkoth, J., Vitto, L., and Wahl, G. 1981. The cloning and reintroduction into animal cells of a functional CAD gene, a dominant amplifiable genetic marker. *Cell* 27:267-277.

Doms, R.W., Lamb, R.A., Rose, J.K., and Helinius, A. 1993. Folding and assembly of viral membranes. *Virology* 193:545-562.

Eagle, H. 1955. Nutritional needs of mammalian cells in tissue culture. *Science* 122:17-20.

Ellis, R.J. and Hemmingsen, S.M. 1989. Molecular chaperones: Proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem. Sci.* 14:339-342.

Gawlitzek, M., Conradt, H.S., and Wagner, R. 1995. Effect of different cell culture conditions on the polypeptide integrity and N-glycosylation of a recombinant glycoprotein. *Biotechnol. Bioeng.* 46:536-544.

Gebert, C.A. and Gray, P.P. 1995. Expression of FSH in CHO cells. 2. Stimulation of hFSH expression levels by defined medium supplements. *Cytotechnology* 17:13-19.

Gething, M.J. and Sambrook, J. 1992. Protein folding in the cell. *Nature* 335:33-45.

BOSTON UNIVERSITY CENTRAL LIBRARY

Goochee, C.F., Gramer, M.J., Andersen, D.C., Bahr, J.B., and Rasmussen, J.R. 1991. The oligosaccharides of glycoproteins: Bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties. *Bio/Technology* 9:1347-1355.

Gramer, M.J. and Goochee, C.F. 1993. Glycosidase activities in Chinese hamster ovary cell lysate and cell culture supernatant. *Biotechnol. Prog.* 9:366-377.

Gramer, M.J., Goochee, C.F., Chock, V.J., Brousseau, D.T., and Sliwkowski, M.B. 1995. Removal of sialic acid from a glycoprotein in CHO cell culture supernatant by action of an extracellular CHO cell sialidase. *Bio/Technology* 13:692-698.

Gray, D.R., Chen, S., Howarth, W., Inlow, D., and Maiorella, B.L. 1996. CO_2 in large-scale and high-density CHO perfusion culture. *Cytotechnology* 22:65-78.

Gu, M.B., Todd, P., and Kompala, D.S. 1993. Foreign gene expression (β-galactosidase) during the cell cycle phases in recombinant CHO cells. *Biotechnol. Bioeng.* 42:1113-1123.

Gu, M.B., Todd, P., and Kompala, D.S. 1996. Cell cycle analysis of foreign gene (β-galactosidase) expression in recombinant mouse cells under control of mouse mammary tumor virus promoter. *Biotechnol. Bioeng.* 50:229-237.

Han, K.K. and Martinage, A. 1992. Post-translational chemical modification(s) of proteins. *Int. J. Biochem.* 24:19-28.

Hart, G.W., Holt, G.H., and Haltiwanger, R.S. 1988. Nuclear and cytosolic glycosylation. *Trends Biochem. Sci.* 13:380-384.

Hassel, T. and Butler, M. 1990. Adaptation to non-ammoniagenic medium and selective substrate feeding lead to enhanced yields in animal cell cultures. *J. Cell Sci.* 96:501-508.

Hayter, P.M., Curling, E.M., Baines, A.J., Jenkins, N., Salmon, I., Strange, P.G., Tong, J.M., and Bull, A.T. 1992. Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-γ. *Biotechnol. Bioeng.* 39:327-335.

Hayter, P.M., Curling, M.A.E., Gould, M.L., Baines, A.J., Jenkins, N., Salmon, I., Strange, P.G., and Bull, A.T. 1993. The effect of dilution rate on CHO cell physiology and recombinant interferon-gamma production in glucose limited chemostat culture. *Biotechnol. Bioeng.* 42:1077-1085.

Hendershot, L.M., Ting, J., and Lee, A.S. 1988. Identity of the immunoglobulin heavy-chain-binding protein with the 78000 dalton glucose-regulated protein and the role of post translational modifications in its binding function. *Mol. Cell. Biol.* 8:4250-4256.

Hickman, S. and Kornfeld, S. 1978. Effect of tunicamycin on IgM, IgA, and IgG secretion by mouse plasmacytoma cells. *J. Immunol.* 121:990-996.

Hilton, D.J., Watowich, S.S., Murray, P.J., and Lodish, H.F. 1995. Increased cell surface expression and enhanced folding in the endoplasmic reticulum of a mutant erythropoietin receptor. *Proc. Natl. Acad. Sci. U.S.A.* 92:190-194.

Hirschberg, C.B. and Snider, M.D. 1987. Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annu. Rev. Biochem.* 56:63-89.

Hooker, A.D., Goldman, M.H., Markham, N.H., James, D.C., Ison, A.P., Bull, A.T., Strange, P.G., Salmon, I., Baines, A.J., and Jenkins, N. 1995. N-glycans of recombinant human interferon-gamma change during batch culture of Chinese hamster ovary cells. *Biotechnol. Bioeng.* 48:639-648.

Huovila, A-P., Eder, A.M., and Fuller, S.D. 1992. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. *J. Cell Biol.* 118:1305-1320.

Hurt, E.C., Peshold-Hurt, B., and Schatz, G. 1984. The cleavable prepiece of an imported mitochondrial protein is sufficient to direct cytosolic dihydrofolate reductase into the mitochondrial matrix. *FEBS Lett.* 178:306-310.

Hurtley, S.M. and Helenius, A. 1989. Protein oligomerization in the endoplasmic reticulum. *Annu. Rev. Cell Biol.* 5:277-307.

Hwang, C., Sinskey, A.J., and Lodish, H.F. 1992. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* 257:1496-1502.

Jackson, M.R., Nilsonn, T., and Peterson, P.A. 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* 9:3153-3162.

Jayme, D.W. 1991. Nutrient optimization for high density biological production operations. *Cytotechnology* 5:15-30.

Jenkins, N. 1990. Growth factors. *In* Mammalian Cell Biotechnology (M. Butler, ed.) pp. 39-55. IRL Press, Oxford.

Jenkins, N., Castro, P.M.L., Menon, S., Ison, A.P., and Bull, A.T. 1994. Effect of lipid supplements on the production and glycosylation of recombinant interferon-γ expressed in CHO cells. *Cytotechnology* 15:209-215.

Jenkins, N., Parekh, R.B., and James, D.C. 1996. Getting the glycosylation right: Implications for the biotechnology industry. *Nature Biotechnol.* 14:975-981.

Kane, S.E., Reinhard, D.H., Fordis, C.M., Pastan, I., and Gottesman, M.M. 1989. A new vector using the human multidrug resistance gene as a selectable marker enables overexpression of foreign genes in eukaryotic cells. *Gene* 84:439-446.

Kaufman, R.J. 1990a. Use of recombinant DNA technology for engineering mammalian cells to produce proteins. *In* Large-Scale Mammalian Cell Culture Technology (A.J. Lubiniecki, ed.) pp. 15-69. Marcel Dekker, New York.

Kaufman, R.J. 1990b. Selection and coamplification of heterologous genes in mammalian cells. *Methods Enzymol.* 185:537-566.

Kaufman, R.J. and Sharp, P.A. 1982. Amplification and expression of sequences cotransfected with

RCUIAS LIBRARY UNIVERSITY LIBRARY

a modular dihydrofolate reductase complimentary DNA gene. *J. Mol. Biol.* 159:601-621.

Kaufman, R.J., Wasley, L.C., Furie, B.C., Furie, B., and Schoemaker, C. 1986a. Expression, purification and characterization of recombinant γ-carboxylated factor IX synthesized in Chinese hamster ovary cells. *J. Biol. Chem.* 261:9622-9628.

Kaufman, R.J., Murtha, P., Ingolia, D.E., Yeung, C.Y., and Kellems, E.R. 1986b. Selection and amplification of heterologous genes encoding adenosine deaminase in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 83:3136-3140.

Keenan, J. and Clynes, M. 1996. Replacement of transferrin by simple iron compounds for MDCK cells grown and subcultured in serum-free medium. [Letter]. *In Vitro Cell. Dev. Biol. Anim.* 32:451-453.

Keller, G.-A., Gould, S., Deluca, M., and Subramani, S. 1987. Firefly luciferase is targeted to peroxisomes in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 84:3264-3268.

Khan, M.W., Musgrave, S.C., and Jenkins, N. 1995. N-linked glycosylation of tissue plasminogen activator in Namalwa cells. *Biochem. Soc. Trans.* 23:S99.

Kingsley, D.M., Kozarsky, K.F., Hobbie, L., and Krieger, M. 1986. Reversible defects in O-linked glycosylation and LDL-receptor expression in UDP-Gal/UDP-GalNAc 4-epimerase deficient mutant. *Cell* 44:749-759.

Kingston, R.E., Kaufman, R.J., Beddington, C.R., and Rolfe, M.R. 1993. Amplification using CHO cell expression vectors. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.14.1-16.14.13. John Wiley & Sons, New York.

Klinman, D.M. and McKearn, T.J. 1981. Dialyzable serum components can support the growth of hybridoma cell lines in vitro. *J. Immunol. Methods* 42:1-9.

Kornfeld, R. and Kornfeld, S. 1985. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54:631-664.

Lee, E.U., Roth, J., and Paulson, J.C. 1989. Alteration of terminal glycosylation sequences on N-linked oligosaccharides of Chinese hamster ovary cells by expression of β-galactosidase α2,6-sialyltransferase. *J. Biol. Chem.* 264:981-1008.

Leelavatcharamas, V., Emery, A.N., and Al-Rubeai, M. 1994. Growth and interferon-gamma production in batch culture of CHO cells. *Cytotechnology* 15:65-71.

Lifely, M.R., Hale, C., Boyce, S., Keen, M.J., Phillips, J. 1995. Glycosylation and biological activity of CAMPATH-1H expressed in different cell lines and grown under different culture conditions. *Glycobiology* 5:813-822.

Lodish, H.F. and Kong, N. 1990. Perturbation of calcium blocks exit of secretory proteins from the rough endoplasmic reticulum. *J. Biol. Chem.* 265:10893-10899.

Machamer, C.E. and Rose, J.K. 1987. A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *J. Cell Biol.* 105:1205-1214.

Malim, M.H., Boehnlein, S., Hauber, J., and Cullen, B.R. 1989. Functional dissection of the HIV-1 Rev trans-activator-derivation of a trans-dominant repressor of Rev function. *Cell* 33:153-159.

Mariani, B.D., Slate, D.L., and Schimke, R.T. 1981. S phase–specific synthesis of dihydrofolate reductase in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. U.S.A.* 78:4985-4989.

Martinelle, K., Westlund, A., and Haggstrom, L. 1996. Ammonium ion transport—a cause of cell death. *Cytotechnology* 22:251-254.

Mayer, T., Tamura, T., Falk, M., and Niemann, H. 1988. Membrane integration and intracellular transport of coronavirus glycoprotein E1, a class III membrane glycoprotein. *J. Biol. Chem.* 263:14956-14963.

McIlhinney, R.A.J. 1990. The fats of life: The importance and function of protein acylation. *Trends Biochem. Sci.* 15:387-391.

McKnight, S. and Tijan, R. 1986. Transcriptional selectivity of viral genes in mammalian cells. *Cell* 46:795-805.

McQueen, A. and Bailey, J.E. 1990. Effect of ammonium ion and extracellular pH on hybridoma cell metabolism and antibody production. *Biotechnol. Bioeng.* 35:1067-1077.

Miyazawa, S., Osumi, T., Hashimoto, T., Ohyno, K., Miura, S., and Fujiki, Y. 1989. Peroxisome targetting signal of rat liver acyl-coenzyme A oxidase resides in the carboxy terminus. *Mol. Cell. Biol.* 9:83-91.

Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P., and Chambon, P. 1981. The SV-40 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucl. Acids Res.* 9:6047-6059.

Mortensen, R., Chesnut, J.D., Hoeffler, J.P., and Kingston, R.E. 1997. Selection of transfected mammalian cells. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 9.5.1-9.5.19. John Wiley & Sons, New York.

Moss, B. and Earl, P.L. 1991. Overview of the vaccinia virus expression system. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.15.1-16.15.5. John Wiley & Sons, New York.

Munro, S. and Pelham, H.R.B. 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48:899-907.

Murhammer, D. and Goochee, C.F. 1990. Sparged animal cell bioreactors: Mechanism of cell damage and Pluronic F68 protection. *Biotechnol. Prog.* 6:391-397.

Myers, M.G., Jr., Sun, X.J., and White, M.F. 1994. The IRS-1 signaling system. *Trends Biochem. Sci.* 19:289-293.

ROYAL / CENTRAL / FIRST LIBRARY

Neuhaus, G., Neuhaus-Url, G., Gruss, P., and Schweiger, H.G. 1984. Enhancer-controlled expression of the simian virus 40 T-antigen in the green alga *Acetabularia*. *EMBO J.* 3:2169-2172.

Neurath, H. 1989. Proteolytic processing and physiological regulation. *Trends Biochem. Sci.* 14:268-271.

Okamoto, M., Nakai, M., Nakayama, C., Yanagi, H., Matsui, H., Noguchi, H., Namiki, M., Sakai, J., Kadota, K., and Fukui, M. 1991. Purification and characterization of three forms of differently glycosylated recombinant human granulocyte macrophage colony stimulating factor. *Arch. Biochem. Biophys.* 286:562-568.

Page, M.J. and Sydenham, M.A. 1991. High level expression of the humanized monoclonal antibody CAMPATH-1H in Chinese hamster ovary cells. *Biotechnology* 9:64-68.

Pak, S.C.O., Hunt, M.W., Bridges, M.J., Sleigh, M.J., and Gray, P.P. 1996. Super-CHO: A cell line capable of autocrine growth under fully defined protein-free conditions. *Cytotechnology* 22:139-146.

Pelham, H.R.B. and Bienz, M. 1982. A synthetic heat-shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene. *EMBO J.* 1:1473-1477.

Pouwels, P.H., Enger-Valk, B.E., and Brammar, W.J. (eds.). 1988. Vectors for animal cells. *In* Cloning Vectors (ch. VIII-I). Elsevier, Amsterdam.

Reitman, M.L. and Kornberg, S. 1981. Lysosomal enzyme targetting: N-acetylglucosaminyl-phosphotransferase selectively phosphorylates native lysosomal enzymes. *J. Biol. Chem.* 256:11977-11980.

Rosenquist, G.L. and Nicholas, H.B., Jr. 1993. Analysis of sequence requirements for protein tyrosine sulfation. *Protein Sci.* 2:215-222.

Rothman, J.E. and Orci, L. 1992. Molecular dissection of the secretory pathway. *Nature* 335:409-415.

Rothman, M.E. 1994. Mechanisms of intracellular protein transport. *Nature* 372:55-63.

Shelikoff, M., Sinskey, A.J., and Stephanopoulos, G. 1994. The effect of protein synthesis inhibitors on the glycosylation site occupancy of recombinant human prolactin. *Cytotechnology* 15:195-208.

Shitara, K., Nakamura, K., Tokutake-Tanaka, Y., Fukushima, M., and Hanai, N. 1994. A new vector for the high level expression of chimeric antibodies in myeloma cells. *J. Immunol. Methods* 167:271-278.

Simonsen, C. and Levinson, A. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 80:2495-2499.

Stanley, P. 1989. Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell. Biol.* 9:377-383.

Stanley, P. 1991. Glycosylation engineering: CHO mutants for the production of glycoproteins with tailored carbohydrates. *In* Protein Glycosylation: Cellular, Biotechnological and Analytical Aspects (H.S. Conradt, ed.) pp. 225-234. GBF Monographs, VCH Publishers, New York.

Stenflo, J., Holme, E., Lindstedt, S., Chandramuli, N., Huang, L., Tam, J., and Merrifield, R. 1989. Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 86:444-447.

Suttie, J.W. 1985. Vitamin K–dependent carboxylase. *Annu. Rev. Biochem.* 54:459-477.

Teige, M., Weidermann, R., and Kretzmer, G. 1994. Problems with serum-free production of antithrombin III regarding proteolytic activity and product quality. *J. Biotechnol.* 34:101-105.

Tsang, T.C., Harris, D.T., Akporiaye, E.T., Chu, R.S., Brailey, J., Liu, F., Vasanwala, F.H., Schluter, S.F., and Hersch, E.M. 1997. Mammalian expression vector with two multiple cloning sites for expression of two foreign genes. *BioTechniques* 22:68.

Urlaub, G. and Chasin, L.A. 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U.S.A.* 77:4216-4220.

Von Heijne, G. 1983. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133:17-21.

Von Heijne, G. 1984. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: Implications for protein export in prokaryotic and eukaryotic cells. *EMBO J.* 3:2315-2318.

Walter, P. and Lingappa, V.R. 1986. Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell. Biol.* 2:499-516.

Wirth, M. and Hauser, H. 1993. Genetic engineering of animal cells. *In* Biotechnology: A Multivolume Comprehensive Treatise, Vol. 2, 2nd ed., pp. 663-744. VCH Publishers, New York.

Wurm, F.M and Petropoulos, J. 1994. Plasmid integration, amplification and cytogenetics in CHO cells: Questions and comments. *Biologicals* 22:95-102.

Yeung, C., Ingolia, E., Bobonis, C., Dunbar, B., Riser, M., Siciliano, M., and Kellems, R. 1983. Selective overproduction of adenosine deaminase in cultured mouse cells. *J. Biol. Chem.* 258:8338-8346.

Contributed by David Gray
Chiron Corporation
Emeryville, California

RCU IAS CONTROL UNIVERSITY LIBRARY

Production of Recombinant Proteins in Mammalian Cells

Mammalian cells are the relevant hosts for expression of properly folded glycoproteins that possess post-translational modifications and that have biological activities similar to those of proteins derived from human cells (*UNIT 5.9*). Increasing interest in the production of glycoproteins (Chapter 12) for clinical application and as reagents for the study of biological mechanisms has led to greater use of mammalian cell protein expression systems. In response to this, developments in molecular biology and process science have provided mammalian cell expression vectors, cell lines, and growth media—and have established methods for high-density suspension culture in compact bioreactors.

When DNA is transferred into a mammalian host cell, the genes may be expressed without being integrated into the chromosomal DNA (transient expression). A small percentage of cells incorporate the DNA into a transcriptionally active chromosomal locus and may express the gene on a permanent basis from generation to generation (stable expression). In *UNIT 5.9* it was stated that the strategy for expression is dependent on the ultimate goal of the project. When one-time production of a few hundred micrograms of purified protein is sufficient, then rapid, transient ("burst production") expression of the gene in COS-7 cells (Aruffo, 1997), the vaccinia system (Earl and Moss, 1998; Earl et al., 1998a,b; Elroy-Stein and Moss, 1998; Moss and Earl, 1998), Sindbis virus (Invitrogen), or Semliki Forest virus (Liljestrom and Garoff, 1995) are appropriate. The best strategy for consistent production of larger quantities of pure protein is stable expression. Popular hosts for stable expression are Chinese hamster ovary (CHO) cells, baby hamster kidney (BHK-21) cells, myeloma cells, and the transformed kidney cell line 293. Protocols for stable production in CHO cells are described in this unit. Typical methods for transfection using commercially available plasmid expression vectors (see Table 5.10.1) are described, along with methods to select for stable expression and methods for amplifying the expression level in the transfected cell. Following this, procedures are presented for efficient cell growth to obtain significant amounts of protein product. A general flow scheme of the overall process is provided in Figure 5.9.1.

Stable expression is obtained when the gene of interest and associated expression elements (*cis* elements) are incorporated into a transcriptionally active region of the host cell chromosome. Production of the plasmid vector for use in transfection is usually performed in *E. coli* cells and the plasmid is purified using standard column chromatography methods (see Basic Protocol 1). DNA is transferred into the cell by lipofection (see Basic Protocol 2) or electroporation (see Alternate Protocol 1). Derivation of stable transfectants requires selection of clones producing the recombinant protein (see Basic Protocol 3, Support Protocol 1, and Alternate Protocol 2) and single-cell cloning (see Support Protocol 2), and as such is a more laborious process than the repeated transient-expression approach. Significant effort and time is needed for several rounds of gene amplification and selection because cell cloning and selection is required at each round. Although the additional work is a possible deterrent to use of this approach, it is possible to produce larger quantities of protein in this way. The other simple advantage of stable transfection is that cell banks can be produced, which enable reproducible production of the product. These advantages will more than provide a dividend after the initial investment of time. To obtain high-level expression, the gene amplification and selection processes can be combined (see Basic Protocol 4).

Growth of the production cells in spinner flasks and/or bioreactors is necessary to obtain sufficient total protein for further purification and study. Procedures for adapting cells to

Contributed by Su Chen, David Gray, Johnny Ma, and Shyamsundar Subramanian
Current Protocols in Protein Science (1998) 5.10.1-5.10.41
Copyright © 1998 by John Wiley & Sons, Inc.

BCIT/CENTRAL UNIVERSITY LIBRARY

Table 5.10.1 Useful Commercially Available Vectors for Transfection of CHO Cells[a]

Strategy	Promoter	Poly(A)	Selectable marker	*E. coli* selectable maker	Plasmids and suppliers	Notes
Dual-plasmid vector cotransfection	SV40, hCMV	SV40	*Neo*, *Pac*	*Amp*	pSI, pCI (Promega); pPUR, pSV2neo (Clontech); pSVK3 (Pharmacia),	
Single-plasmid vector (regular cassette)	hCMV, RSV, SV40, MMTV	SV40, BGH	*Neo*, *Zeo*	*Amp*, *Zeo*	pcDNA3.1, pZeoSV2, pRc/CMV2 , pRc/RSV (Invitrogen); pMAMneo (Clontech); pCI-neo (Promega),	
	hCMV	SV40	*GS*	*Amp*	pEE14 (CellTech)	Amplification using glutamine synthetase with MSX
Single plasmid vector (dicistronic cassette)	hCMV	BGH	*Neo*, *Hyg*	*Amp*	pIRES1neo, pIRES1hyg (Clontech)	Uses encephalomyo-carditis virus IRES sequence
Special plasmid vectors	pCMV	BGH	*Zeo*	*Amp*	pSecTag (Invitrogen)	Secretion of proteins lacking signal peptides
					pHook (Invitrogen)	Isolation of positive clones through magnetic affinity cell sorting (MACS)
					pTracer-CMV (Invitrogen)	Isolation of positive clones using GFP and FACS

[a]Abbreviations: hCMV, human cytomegalovirus (CMV) major immediate early promoter; *Hyg*, hygromycin phosphotransferase B gene; MSX, methionine sulfoximine gene; *Neo*, aminoglycoside phosphotransferase II gene; *Pac*, puromycin resistance gene; RSV, Rous sarcoma virus long terminal repeat; *Zeo*, Zeocin resistance gene.

suspension conditions and low-serum media are included (see Basic Protocol 5) as well as procedures for growing cells in various vessels and in batch culture (see Basic Protocols 6 and 7), and in continuous and perfusion cultures (see Alternate Protocol 3). Most often, the product is secreted into the medium and must be separated from the cells prior to subsequent storage or purification (see Basic Protocol 8). In some cases, the protein of interest may be intracellular or membrane-associated, and thus the cells become the product and must be harvested from the culture (see Basic Protocol 9).

Support protocols describe freezing of cells (see Support Protocol 3), determination of growth rates (see Support Protocols 4 and 5), determination of specific productivity of cells (see Support Protocols 6 and 7), preparing samples for assay (see Support Protocol 8), and setting up a 10-day shaker-flask growth curve (see Support Protocol 9).

NOTE: Milli-Q purified water or equivalent is used for all aqueous buffer reagents.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO_2 incubator unless otherwise noted.

PCU IASRI CENTRAL UNIVERSITY LIBRARY

PLASMID PURIFICATION BY ALKALINE-LYSIS/ANION-EXCHANGE CAPTURE

The gene of interest (in the form of cDNA) is inserted into the multiple cloning region of the plasmid (which has been constructed with a number of restriction sites) after any required modifications in the sequence. Commercially available plasmids provide various choices of elements; a subset are shown in Figure 5.10.1.

Sufficient purified plasmid is produced in competent *E. coli* cells (e.g., strain DH5α) using the preparation approach discussed here. The alkaline-lysis/anion-exchange capture method is commonly used to prepare large amounts of DNA for transfection of host cells. A number of commercially available kits are available, as mentioned below, making this method quick and easy to use. Following DNA preparation, transfection of the CHO cells is performed using lipofection (see Basic Protocol 2) or electroporation (see Alternate Protocol 1).

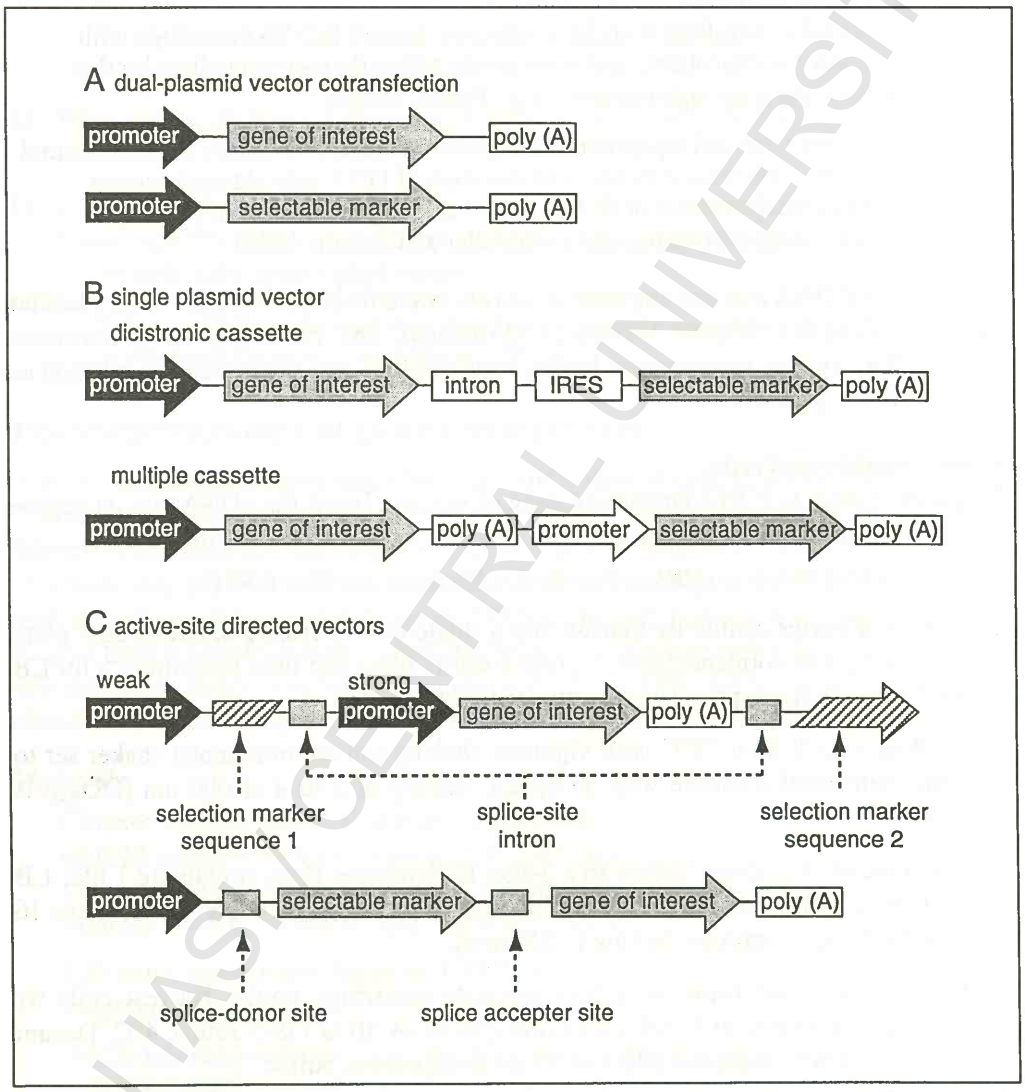


Figure 5.10.1 Expression vector elements for common transfection strategies. (**A**) Dual- and (**B**) single-plasmid vectors are shown for commercial vectors types described in Table 5.10.1. The active site–directed vectors (**C**) are reported throughout the literature (Lucas, 1996; Wurm et al., 1996); although not currently available as commercial catalog items, they may provide advantages in obtaining high expression levels. Abbreviations: IRES, internal ribosomal entry site.

BIAS / CENTRAL UNIVERSITY LIBRARY

Materials

cDNA containing gene of interest
Plasmid with appropriate *cis*-acting elements (Fig. 5.10.1 and *UNIT 5.2*)
Competent *E. coli* cells (e.g., strain DH5-α; Life Technologies)
LB medium (*UNIT 5.2*) containing antibiotics for selection (e.g, 100 µg/ml ampicillin, 40 µg/ml kanamycin, or 25 µg/ml tetracycline, depending on plasmid kit; all antibiotics available from Sigma)
Resuspension buffer (see recipe)
Lysis solution: 0.2 M NaOH/1% (w/v) SDS
Neutralization solution: 1.3 to 3.0 M potassium acetate, pH 4.8 to 5.5
Isopropanol
TE buffer (*APPENDIX 2E*)
80% ethanol

Sterile glass test tubes for bacterial culture
Environmental shaker (New Brunswick Scientific)
3-liter Erlenmeyer flask
Beckman J2-21 centrifuge with JA-10 rotor or Sorvall RC-5B centrifuge with GS-3 rotor (or equivalent) and appropriate polycarbonate centrifuge bottles
UV/visible light spectrophotometer (e.g., Perkin-Elmer)

Additional reagents and equipment for ligation of linkers to cDNA (Klickstein and Neve, 1991), plasmid selection and insertion of DNA into plasmid vectors (*UNIT 5.2*), transformation of *E. coli* (*UNIT 5.2*), and purification of DNA by anion-exchange chromatography (Budelier and Schorr, 1998)

NOTE: Several DNA plasmid preparation kits are available commercially—e.g., Plasmid Maxi Kit (Qiagen), Wizard Maxipreps (Promega), and FlexiPrep Kit (Pharmacia Biotech). Kits include resuspension buffer, lysis solution, and neutralization solution as described in the protocol.

Prepare transformed cells

1. Ligate linkers to cDNA containing gene of interest. Insert the cDNA into an appropriate plasmid vector and transform *E. coli* host cells with the vector (*UNIT 5.2*).

   Ligation of linkers to cDNA is described in Klickstein and Neve (1991).

2. Prepare a starter culture by transferring a single colony from a selective agar plate containing transformants (*UNIT 5.2*) into a sterile glass test tube containing 5 ml LB medium with the appropriate selective antibiotic.

3. Incubate 6 to 8 hr at 37°C with vigorous shaking (on environmental shaker set to ~300 rpm) until a culture with an optical density of 2 to 4 at 600 nm (OD_{600}) is obtained.

4. Add 1 ml of this starter culture to a 3-liter Ehrlenmeyer flask containing 1 liter LB medium with the same selective antibiotic used in step 2. Incubate culture 12 to 16 hr at 37°C with vigorous shaking (~300 rpm).

5. Transfer cell suspension to a polycarbonate centrifuge bottle. Harvest cells by centrifuging 15 min at 6,000 × *g* (6,000 rpm in JA-10 or GS-3 rotor), 4°C. Decant supernatant and resuspend pellet in 30 ml resuspension buffer.

Isolate plasmid DNA

6. Add 30 ml of lysis solution to resuspended cells. Mix thoroughly by inverting (not vortexing) and incubate ~10 min at room temperature.

   Cell lysis is complete when solution becomes relatively clear and viscous.

ECU IAS / CENTRAL UNIVERSITY LIBRARY

7. Add 30 ml of neutralization solution to the lysed cells and mix immediately by inverting.

8. Centrifuge 20 min at 10,000 to14,000 × *g*, (7,500 to 9,000 rpm in JA-10 or GS-3 rotor), 4° to 8°C.

   This step separates the cell debris from the lysate. The plasmid will be in the lysate.

9. Precipitate crude DNA from the clear lysate by adding 0.5 vol isopropanol, then centrifuge 30 min at 15,000 × *g* (9,200 rpm in JA-10 or GS-3 rotor), 4° to 8°C.

   Mark the region in the centrifuge tube where the pellet is expected to stick, as it may not be visible after isopropanol precipitation.

10. Dissolve DNA pellet in 2 ml TE buffer.

Purify plasmid DNA

11. Apply resuspended DNA solution to an anion-exchange resin (e.g., as provided with Qiagen Plasmid Maxi Kit; also see Budelier and Schorr, 1998).

    The plasmid DNA will bind to the resin. Budelier and Schorr (1998) describes the anion-exchange procedure for DNA.

12. Wash resin with 80% ethanol or with the wash buffer provided with the kit (Budelier and Schorr, 1998).

13. Elute DNA from resin with TE buffer or the elution buffer provided with the kit. Determine DNA concentration by measuring OD_{260} (50 µg/ml = 1.0 OD_{260}) using a UV/visible light spectrophotometer.

    The eluted DNA can be used directly in transfection of host cells (see Basic Protocol 2 and see Alternate Protocol 1) or stored at –20°C.

TRANSFECTION OF CELLS BY LIPOFECTION

BASIC PROTOCOL 2

A good method for transfecting CHO cells and adherent cells in general is lipofection. This technique can typically achieve a 50% transfection efficiency (Subramanian and Srienc, 1996), although the value depends on the sensitivity of detection. In the process of lipofection, plasmid DNA is complexed with cationic lipids; the lipid then interacts with the cell membrane and delivers the DNA to the cell. In the protocol below, the lipid reagent used is Lipofectamine, which has been very effective with CHO cells. There are several other lipids available commercially from Life Technologies, Boehringer Mannheim, and Invitrogen, which may be more suitable for other cell types.

Materials

Chinese hamster ovary (CHO) cells (Table 5.9.1)
αMEM medium with and without 10% FBS (see recipe)
Lipofectamine (Life Technologies; store up to 6 months at 4°C)
Plasmid DNA to be transfected (see Basic Protocol 1)

100-mm tissue culture dishes or T-75 flasks
Polystyrene tubes

Additional reagents and equipment for growing mammalian cells in culture (*APPENDIX 3C*)

1. Plate CHO cells, at a density of 1.5–3.0 × 10^6 viable cells per 100-mm tissue culture dish or T-75 flask, using αMEM medium/10% FBS. Incubate ~24 hr to achieve 50% 80% confluency (*APPENDIX 3C*).

BOWIE STATE UNIVERSITY LIBRARY

2. Dilute 48 µl of Lipofectamine in 0.8 ml serum-free αMEM in a polystyrene tube. Dilute 8 µg of DNA to be transfected in 0.8 ml serum-free αMEM in a separate polystyrene tube.
3. Mix lipid and DNA solutions and incubate 45 min at room temperature to form complexes. Dilute the complexes in 6.4 ml serum-free αMEM.
4. Wash cells with 10 to 20 ml serum-free αMEM per dish, remove the medium, then add the diluted lipid-DNA complexes to cells.

   Alternatively, the lipid-DNA complexes can be directly diluted in the tissue culture dish.
5. Incubate cells with the complexes for 5 to 6 hr under normal culture conditions. Replace complexes with αMEM/10% FBS, then incubate 24 to 48 hr.

   Note that the recipe for αMEM (see Reagents and Solutions) includes no antibiotics.

   Cells are ready for either transient analysis or selection (see Basic Protocol 3) 24 to 48 hr post-transfection.

ALTERNATE PROTOCOL 1

TRANSFECTION OF CELLS BY ELECTROPORATION

Electroporation allows delivery of DNA into cells by creation of transient pores in the cell membrane. In many cell types that grow in suspension, the lipofection method (see Basic Protocol 2) does not result in high transfection efficiencies. In these situations electroporation is an appropriate choice, as it has been shown to work with numerous cell types (both suspension and adherent). Electroporation may also be the preferred method when single-copy integration is desired, since this technique delivers lower amounts of DNA to the cell than lipofection. Both methods work well in the case of CHO cells. The authors have found that the Bio-Rad Gene Pulser II, equipped with a capacitance extender, is suitable for electroporation of mammalian cells. This instrument delivers a decaying exponential pulse. Other instruments—e.g., from BTX and Invitrogen—can also be used.

Additional Materials (also see Basic Protocol 2)

Electroporator and electroporation cuvettes with 0.4-cm electrode gap (Bio-Rad)

Additional reagents and equipment for growing mammalian cells in culture, trypsinizing cells, and counting viable cells by trypan blue exclusion (*APPENDIX 3C*)

1. Harvest exponentially growing (subconfluent) cultures by trypsinization and count cells (*APPENDIX 3C*). Centrifuge cells 10 min at 200 to 500 × *g*, room temperature. Resuspend pellet in αMEM/10% FBS medium at 1–2 × 10^7 viable cells/ml. Dispense 0.5-ml aliquots into electroporation cuvettes.
2. Add 10 to 20 µg of DNA to be transfected and incubate 10 min on ice.
3. Resuspend cells by flicking, then pulse at high capacitance (~1000 µF) at 200 to 400 V (which should result in a time delay in 5 to 20 msec range). Use 10- to 50-V steps.
4. Incubate electroporated cells 10 min at 37°C. Transfer cells into a T-75 flask or 100-mm dish with 20 ml of medium and incubate 24 to 48 hr under normal culture conditions (*APPENDIX 3C*).

   The cuvette may need to be rinsed for complete removal of cells.

   Cells are ready for either transient analysis or selection (see Basic Protocol 3) 24 to 48 hr post-transfection.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

SELECTION FOR NEOMYCIN PHOSPHOTRANSFERASE (NPTII) WITH GENETICIN (G418)

Following the transfection, cells are subjected to the selection process, in which they are exposed to a drug that allows cells expressing the resistance marker to survive while killing the rest. The commonly used dominant selection markers include neomycin phosphotransferase (NPTII, aminoglycoside phosphotransferase II) which enables growth in the presence of G418 (geneticin). Other selection approaches in common use are the glutamine synthetase (GS) system, which uses growth in glutamine-free medium for selection (Kingston et al., 1993). Another alternative and common approach is the use of a recessive marker such as dihydrofolate reductase (DHFR) with initial selection in nucleoside-free medium (see Basic Protocol 4).

Selection, as described in this protocol, and screening (see Alternate Protocol 2), are the two different methods that help isolate cells that express the marker gene. The marker gene is an indirect indication that the protein of interest is also being expressed. In selection, the drug kills cells that do not express the marker. This protocol describes selection with the commonly used dominant selection marker neomycin phosphotransferase II.

Materials

Culture of transformed cells at 50% to 70% confluency, in T-25 or T-75 flasks (24 to 48 hr post-transfection; see Basic Protocol 2 or Alternate Protocol 1)

Selective medium: αMEM medium (see recipe) containing 5% to 10% FBS and 0.1 to 1.0 mg/ml G418 (Geneticin, Life Technologies; add from 50 mg/ml stock prepared in 100 mM HEPES, pH 7.0 to 7.2; see Support Protocol 1 for optimization of concentration)

96-well flat-bottom tissue culture plates

Inverted microscope

T-75 tissue culture flasks

Additional reagents and equipment for growing mammalian cells in culture, trypsinizing cells, and counting viable cells by trypan blue exclusion (*APPENDIX 3C*)

1. At 24 to 48 hr post-transfection, harvest cells by trypsinization and count viable cells by trypan blue exclusion (*APPENDIX 3C*).
2. Plate 100 to 1000 cells in each well of a 96-well tissue culture plate using 200 µl selective medium per well. Begin incubation.
3. Replace selective medium once every 3 to 4 days and observe cells using an inverted microscope to monitor cell death and detachment.

   Large colonies of live cells (containing 100 to 200 cells) should appear within 3 weeks.
4. As the colonies grow and become confluent (indicated by change in color of the phenol red indicator in the medium from dark red to orange as the pH decreases), screen individual clones for the protein of interest using different methods depending on the targeting of the gene product.

   The supernatant is assayed for secreted proteins using an ELISA. For cytoplasmic proteins, the cells are screened through the detection of surrogate markers such as green fluorescent protein (GFP; Kahana and Silver, 1996). Membrane proteins may be amenable to immunofluorescent staining followed by flow cytometric sorting. An alternative approach is to pick 10 to 20 clones and screen cell lysates with ELISAs or immunoblots (UNIT 10.8). If the protein is localized to the membrane protein, immunofluorescent staining with specific antibodies (UNIT 10.10) can be directly applied.

BCU IAS/CENTRAL UNIVERSITY LIBRARY

5. Trypsinize viable productive colonies and transfer to wells of a 24-well plate, each containing 2 ml selective medium. After cells reach confluency, trypsinize and transfer to a T-75 flask with 20-ml medium (*APPENDIX 3C*).

When the attachment culture in a tissue culture flask grows to 80% to 90% confluency, it should be trypsinized and subcultured into a new flask with a starting cell density of 1 × 10^4/cm^2 for continuation of cell growth.

SUPPORT PROTOCOL 1

DETERMINATION OF BACKGROUND SENSITIVITY TO G418

The drug concentration and the plating cell density are very important for selection to be effective. In the case of selection, the optimal drug concentration is determined as described below. This concentration is set to the minimum required to kill nontransfected cells in 2 to 3 weeks. The selection becomes less stringent when cells are confluent, especially with G418 selection, probably because of lowered metabolic activity. It is important to choose a concentration of the toxic drug that will kill cells that are not expressing the resistance gene but that will allow growth.

Additional Materials (also see Basic Protocol 3)

Selective media containing various concentrations (100 to 1000 μg/ml in 100-μg/ml increments) of G418 (add from 50 mg/ml stock prepared in 100 mM HEPES, pH 7.0 to 7.2)
100-mm tissue culture dishes

1. Plate cells at a density of 0.6–1.2 × 10^6 per 100-mm dish in selective media containing G418 at various concentrations. Begin incubation.
2. Replace medium every 3 to 4 days and observe cells under inverted microscope.
3. Determine the minimum concentration of G418 required for complete killing of cells after 2 weeks.

ALTERNATE PROTOCOL 2

SCREENING OF GREEN FLUORESCENT PROTEIN BY FLOW CYTOMETRY

An alternative to selection (see Basic Protocol 3) is the process of screening, where cells expressing the protein of interest are detected in the population using a noninvasive method and then isolated using flow cytometry or other techniques such as affinity chromatography and magnetic-affinity cell sorting. These methods have not been developed extensively because few noninvasive detection methods have become available. In the protocol below, the green fluorescent protein (GFP; Kahana and Silver, 1996) has been used as a screening marker and provides a flow cytometric method for measuring single-cell fluorescence, which may be combined with sorting based on the signal (Subramanian and Srienc, 1996). Any membrane protein against which specific antibodies can be directed may be used instead of GFP. An additional staining step is required when membrane proteins are used.

Materials

Cells transformed with plasmid (see Basic Protocol 2 or Alternate Protocol 1) containing GFP reporter gene (pIRES1-EGFP from Clontech or pTracer-CMV from Invitrogen) at 50% to 70% confluency in T-25 or T-75 flasks
1% (w/v) bovine serum albumin (BSA; Sigma) in PBS (*APPENDIX 2E*)
10× (0.05 mg/ml) propidium iodide stock (PI; Molecular Probes; optional)
Appropriate culture medium

PCU IASTRAL UNIVERSITY LIBRARY

Flow cytometer with cell sorting capability (FACS Vantage and FACSort from Becton Dickinson Immunocytometry Systems or EPICS Elite ESP cell sorter from Coulter Electronics or equivalent)
T-25 tissue culture flasks

Additional reagents and equipment for trypsinizing cells, counting viable cells by trypan blue exclusion, and growing mammalian cells in culture (*APPENDIX 3C*) and flow cytometry (Robinson et al., 1998)

1. At 24 to 48 hr post-transfection, harvest cells by trypsinization and count viable cells by trypan blue exclusion (*APPENDIX 3C*).

   Make sure cells are in single-cell suspension and not in clumps.

2. Resuspend cells at a concentration of ~1×10^6 cells/ml in 1% BSA/PBS. If desired, immediately before analysis, counterstain by adding 0.05 mg/ml (10×) PI to a final concentration of 5 µg/ml.

3. Run cells through flow cytometer at a flow rate of 200 to 1000 cells/sec. Tune laser wavelength to 488 nm and collect the following parameters:

   Forward-angle light scatter (FALS; linear)
   Right-angle light scatter (RALS; linear)
   Green fluorescence: linear and logarithmic using 510- to 540-nm band-pass (FITC) filter
   Red fluorescence (if counterstained with PI): linear or logarithmic using 670 or 690 long-pass filter.

   Gate on viable cells from bivariate plot of FALS versus RALS or FALS versus PI. Set gate to distinguish nonfluorescent from fluorescent cells, then select top 1% to 10% (or all) of the fluorescent cells for sorting (see Robinson et al., 1998, for detailed protocols in flow cytometry and cell sorting).

4. Sort 1×10^3 to 1×10^5 viable cells within the defined fluorescence threshold into a tube containing culture medium. Inoculate cells into T-25 flask with 5 to 10 ml culture medium. Begin incubation.

5. Sort cells again after 2 to 3 weeks, directly into 96-well plates containing 200 µl medium per well, at 1 cell/well. Continue incubation.

   Allow enough time for cells to establish a stable phenotype. Stringent criteria on GFP may be used at this step—i.e., selecting the top 1% or less of fluorescent cells (0.01% to 0.1% of cells can be sorted).

   After the clones grow out they can be screened for the production of the protein of interest.

SINGLE-CELL CLONING

SUPPORT PROTOCOL 2

After cells expressing the gene product have been selected, single-cell cloning is usually applied to obtained monoclonal cultures. This is usually done by diluting a cell suspension and plating the cells in 96-well plates so that only single cells result in colonies, which may then be isolated. The cell suspension should consist mainly of single cells, with no clumps, and the viable cell count should be accurate. Cell clumps after trypsinization should be carefully broken up by repeated pipetting. Since hemacytometer-based counts have a 20% standard deviation (even with two hemacytometers) it is recommended that cells be counted using a Coulter Counter or an Elzone particle counter and that the count for viability (determined by trypan blue exclusion and a hemacytometer; *APPENDIX 3C* be corrected on the basis of the automated cell count to determine the viable cell density).

B. C. U. „M... IAŞI

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

Materials

Cells expressing gene product of interest (see Basic Protocol 3 or Alternate Protocol 1) growing in tissue culture flasks or dishes at 70% to 90% confluency
Appropriate culture medium

Coulter Counter or equivalent
96-well tissue culture plates

Additional reagents and equipment for growing mammalian cells in culture, trypsinizing cells, and counting viable cells by trypan blue exclusion (*APPENDIX 3C*)

1. Trypsinize cells (*APPENDIX 3C*). Break up any cell clumps by pipetting up and down repeatedly, then count cells using a Coulter Counter and determine number of viable cells by trypan blue exclusion.

2a. *Cloning method 1:* Resuspend cells in fresh medium at a concentration such that there will be 0.3 to 1.0 cell per well of a 96-well plate, in 200 to 250 µl of medium per well. Pipet a 200- to 250-µl aliquot of this suspension into each well.

2b. *Cloning method 2 (visual check):* Resuspend in medium at a concentration of 1 cell per 5 to 10 µl medium. Place a 5- to 10-µl drop (a Pasteur pipet can be used for this purpose) in center of flat-bottom 96-well tissue culture plate. Observe ~5 wells at time under an inverted microscope at 100× to 200× magnification, then add 200 to 250 µl of medium to wells with only a single cell.

3. Incubate, replacing medium once every 1 to 2 weeks.

Large colonies with at least 2000 cells take ~2 to 4 weeks to form.

BASIC PROTOCOL 4

AMPLIFICATION OF DIHYDROFOLATE REDUCTASE (DHFR) WITH METHOTREXATE (MTX)

The level of expression after selection is usually not sufficient when the goal is to create a cell line for the production of the protein on a large scale. Amplification to increase gene dosage is a method for creating high producers. Cells are allowed to adapt to growth in the presence of a drug that inhibits the activity of the selection marker. This adaptation usually results from an increase in gene copy number to overcome the inhibition. Stepwise increases in drug concentration and adaptation of the cells amplify the cell line. In this protocol, the selection marker used is dihydrofolate reductase and the drug that stoichiometrically binds and inhibits DHFR activity is methotrexate (MTX). The glutamine synthetase marker and its corresponding inhibiting drug, methionine sulfoximine, is another amplification system that can be used with CHO cells.

Materials

Culture of transformed cells at 50% to 70% confluency, in T-25 or T-75 flasks (24 to 48 hr post-transfection; see Basic Protocol 2 or Alternate Protocol 1)
Nucleoside-free αMEM medium (see recipe but omit nucleosides) containing 10% FBS that has been dialyzed to remove nucleosides (HyClone)
10 mM methotrexate (MTX) stock solution (Sigma; store up to 1 year at –20°C; thaw just before use)

96-well flat-bottom tissue culture plates
24-well tissue culture plates
T-25 tissue culture flasks
100-mm tissue culture dishes

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Additional reagents and equipment for growing mammalian cells in culture, trypsinizing cells, and counting viable cells by trypan blue exclusion (*APPENDIX 3C*)

Perform selection

1. At 24 to 48 hr post-transfection, harvest cells by trypsinization and determine number of viable cells by trypan blue exclusion (*APPENDIX 3C*).
2. Plate 100 to 1000 cells in each well of a 96-well tissue culture plate using 200-µl ml nucleoside-free αMEM/10% dialyzed FBS per well. Begin incubation.
3. As colonies grow and become confluent, screen clones (see Basic Protocol 3, step 4).
4. Trypsinize viable productive colonies and transfer to wells of a 24-well plate, each containing 2 ml nucleoside-free αMEM/10% dialyzed FBS. After cells reach confluency, trypsinize and transfer half to inoculate T-25 flasks containing 5 to 10 ml nucleoside-free αMEM/10% dialyzed FBS. Combine the remaining cells from all clones into one single pool and use in step 5.

   The clones transferred to the T-25 flasks can undergo selection.

   Screening is recommended at this stage before starting the amplification process.

Perform first round of amplification

5. Plate pooled cells at a density of 0.6–1.2 × 10^6 cells/ml in nucleoside-free αMEM/10% dialysed FBS (selective medium) containing 0.02, 0.05, 0.1, and 0.2 µm MTX (added from 10 mM stock) in a series of 100-mm dishes.
6. Replace selective medium in each dish with selective medium containing the same MTX concentration every 3 to 4 days and observe cells using an inverted microscope to monitor cell death and detachment.

   Large colonies of live cells (containing 100 to 200 cells) should appear within 3 weeks.

7. Trypsinize cells and count number of viable cells by trypan blue exclusion (*APPENDIX 3C*).
8. Plate cells from each of the methotrexate concentrations in wells of the same 96-well tissue culture plate at 0.3 to 1.0 viable cells per well in 200 µl selective medium with same MTX concentration as used before. Incubate until confluent, then transfer to 24-well plates and screen clones for productivity (see Basic Protocol 3, step 4).

Perform additional rounds of amplification

9. Plate remaining cells from step 7 at a density of 0.6–1.2 × 10^6 cells per 100-mm dish in selective medium containing an MTX concentration 2- to 5-fold higher than that used before (0.5, 1, 2, 5, and 10 µM MTX).
10. After cells are adapted to the next level of methotrexate carry on with steps 7 to 9.

BCU IAST / CENTRAL UNIVERSITY LIBRARY

ADAPTATION OF SUSPENSION CELLS TO PRODUCTION MEDIUM THROUGH MULTIPLE PASSAGING

Adaptation of transfected CHO cell lines to suspension growth is recommended to enable growth in spinner bottles or reactors. Suspension growth in low-serum medium may improve the initial specific purity of the product by providing the opportunity for higher cell density and product titer than can be obtained by surface-attachment growth, as well as by reducing serum proteins in the medium. Growth of cells in suspension, in medium containing adequate levels of nutrients, enables use of compact production systems in which high cell density can be achieved. Such compact systems are economical choices in laboratory environments.

Materials

Mid–log phase culture of stable-transfected cells in T-75 flask (or thawed 1-ml aliquot of cryopreserved cells; see Support Protocol 2), produced by DHFR selection/MTX amplification (see Basic Protocol 4) or other selection/amplification procedure

Nucleoside-free αMEM medium (see recipe but omit nucleosides) containing 10% FBS that has been dialyzed to remove nucleosides (for DHFR selection/MTX amplification cells) *or* regular αMEM medium (see recipe) containing 10% FBS (prepare as appropriate for cells produced by other selection/amplification procedures)

10 mM methotrexate (MTX) stock solution (Sigma; store up to 1 year at –20°C; thaw just before use)

Dispase solution (Boehringer Mannheim)

DMEM/F12 custom-modified medium (see recipe) with 10%, 1%, and 0.5% dialyzed FBS

CD CHO medium (for DHFR systems; Life Technologies), Super CHO (for non-DHFR systems; Bio-Whittaker), or CHO-S-SFM (for non-DHFR systems; Life Technologies)—optional, for adaptation to serum-free medium

1 g/liter Nucellin-Zn salt (recombinant insulin; Eli Lilly)—optional, for adaptation to serum-free medium

T-175 flasks

Sterile 45-ml polycarbonate screw-cap tubes

Inverted phase-contrast microscope

250-ml shaker flasks

Platform shaker (New Brunswick Scientific)

Additional reagents and equipment for growing mammalian cells in culture, trypsinizing cells, and counting viable cells by trypan blue exclusion (*APPENDIX 3C*), determining specific growth rate of cells (see Support Protocols 4, 5, and 9), determining specific productivity of cells producing heterologous protein (see Support Protocols 6 and 7), and freezing cells (see Support Protocol 2)

Establish cultures in adaptive medium

1. Preincubate a T-175 flask and add 50 ml of prewarmed fresh αMEM/10% FBS (nucleoside-free and containing MTX concentration used for amplification if produced via Basic Protocol 4; prepare as appropriate for other selection/amplification schemes). Keep flask in incubator until the inoculum is prepared.
2. Remove the medium from a mid–log phase (50% to 90% confluent) T-75 flask of stable-transfected cells using a sterile 25-ml pipet. Trypsinize and harvest the cells into a sterile 45-ml polycarbonate tube (*APPENDIX 3C*).

FCU IASI / CENTRAL UNIVERSITY LIBRARY

flask into a shaker flask to begin suspension adaptation, the starting cell density for the new culture should be ~1 × 10^5 viable cells/ml. The procedures for trypsinizing and passaging cells are described in APPENDIX 3C.

3. Determine the concentration of viable cells by trypan blue exclusion (*APPENDIX 3C*), or, if clumping is evident, by adding 0.3 to 2.4 U/ml dispase, incubating 15 min at 37°C, then counting the number of viable cells by trypan blue exclusion. From the viable cell concentration determine the volume of suspension to inoculate into a T-175 flask to obtain a concentration of 1×10^4 viable cells/cm^2. Add the required volume of inoculum to the T-175 flask and incubate.

   CHO cells have a tendency to clump when grown at high cell density in the higher-serum medium. Clumping may be minimized by adapting cells to growth in lower serum (≤0.5% FBS). Reduced-calcium media such as DMEM/F12 also may produce less cell clumping. In cultures that contain clumping cells the only secure counting method is through use of dispase before hemacytometer/trypan blue counting.

4. Periodically observe cells in the tissue culture flask using an inverted phase-contrast microscope at 125× magnification. When the cells reach a confluency of 50%, remove the spent medium and replace with the new medium, DMEM/F12 supplemented with 10% dialyzed FBS and the concentration of MTX (or other selective agent) used previously. Continue incubation, periodically observing growth using the inverted microscope.

5. When cells reach 80% to 95% confluency, trypsinize and harvest the cells. Determine the viable cell density (*APPENDIX 3C*) and calculate the volume of inoculum required to give a final concentration of 1×10^5 viable cells/ml in 100 ml. Simultaneously, prepare a 250-ml shaker flask with 100 ml of fresh DMEM/F12/10% dialyzed FBS/MTX and preincubate to 37°C.

6. Remove the volume of medium calculated for the inoculum from the shaker flask and transfer the corresponding volume of inoculum. Mix by gentle swirling, then remove a 0.25-ml sample and determine post-inoculation viable cell density (*APPENDIX 3C*). Loosely cap the flask, and begin incubation with agitation at 75 to 125 rpm on a platform shaker.

7. Continue incubation and calculate specific growth rate from the growth curve (see Support Protocols 4, 5, and 9) and specific productivity (see Support Protocols 6 and 7). Determine the viable cell density (*APPENDIX 3C*) daily. At 4 days post inoculation, or when the cell density has reached 4×10^5 viable cells/ml, transfer sufficient cells to inoculate a second passage flask as in step 6.

8. Passage the cells in changes of the same medium until the passage-specific growth rate reaches a constant maximum value. At this passage, reduce the serum concentration in the medium to 1% dialyzed FBS.

   Also ascertain that the specific productivity remains within reasonable bounds with increased passage number, as an indication of expression stability.

9. When the passage-specific growth rate shows no further increase, reduce the serum to 0.5% and repeat adaptation passaging.

   When growth rate shows no further gain, the cells have adapted to the low level of serum.

Confirm adaptation to production medium

10. Confirm that cells have adapted to the low level of serum by splitting the cells into two flasks: one with medium plus 10% FBS and another with medium plus 0.5% FBS. Determine growth rate, which show no large difference for correctly adapted cells.

BOLTON CENTRAL UNIVERSITY LIBRARY

11. *If serum-free growth is desired:* Perform a further adaptation from 0.5% to 0% serum by repeating the adaptation steps through a final round using specific-protein-free medium such as CD CHO, Super CHO or CHO-S-SFM (or equivalent). Supplement with 5 mg/liter of Nucellin (added from 1 g/liter stock) and reduce to 1 mg/liter gradually

12. Freeze adapted cells (see Support Protocol 2) or proceed to growth in batch mode (see Basic Protocol 6).

BASIC PROTOCOL 6

GROWTH OF CELLS IN BATCH MODE IN LARGE-SCALE SPINNER CULTURE VESSELS

Suspension-adapted cells may be grown in spinner flasks accommodating liquid volumes of up to 16 liters. Flasks are placed in a humidified temperature-controlled environment, which contains 5% CO_2 gas. The CO_2 is used to provide dissolved carbon dioxide and bicarbonate, which provide nutrient substrate and buffer components in the culture medium. The flasks are set on a rotating platform to agitate the liquid surface and provide adequate oxygen transfer to support cell oxygen demand. The procedure may be initiated from a vial of the cells or from ongoing shaker flasks that are being passaged (see Basic Protocol 5). The ultimate scale of production depends on the expression level, the purification yield, and the quantity of protein required by the researcher. Flasks can accommodate up to 100 ml of culture while small glass spinner jars can be used to produce up to 3 liters of culture. When larger volumes of liquid culture are required, the spinner flasks may become more complicated in design, including dissolved oxygen (dO_2) controllers and pH controllers. Volumes greater than 5 liters of culture can be obtained by using 20-liter spinner jars.

The increased volume and subsequent apparatus size becomes cumbersome and is worthy of arranging in a fixed location. In order to supply adequate oxygen to support the cultures, it is necessary to set up such a spinner jar in a warm room or with an external wraparound band heater (Watlow). The spinner jar is usually fitted with a top-entry mechanical stirrer (Bellco Biotechnology). The supply of oxygen to the culture is achieved through sparging with O_2/air mixtures through a sparge tube fitted with a frit (Mott Cups) and by adding a CO_2/O_2/air mixture into the headspace. Similarly the control of pH is best achieved through use of an in-vessel pH probe with adjustment by addition of 0.5 M NaOH using automatic pH controllers or manually performing daily pH checks. A typical large volume (20-liter) spinner set up is shown in Figure 5.10.2.

Here continued use of MTX is desired, use nucleoside-free medium and dialyzed FBS. In most cases where stable integration and production has been acceptable at levels <1 μM MTX, the continued use of methotrexate should not be required.

Materials

1-ml cryovial or mid–log phase culture of suspension-adapted cells (see Basic Protocol 5)
Medium to which cells have been adapted (see Basic Protocol 5), supplemented with 4 mM L-glutamine and 4.5 to 5.5 g/liter D-glucose
Medical-grade, sterile-filtered 5% CO_2/air mixture
Medical-grade, sterile-filtered 100% O_2 (if necessary)

250-ml screw-cap shaker flasks
Platform shaker (New Brunswick Scientific)
1-, 5-, and 20-liter borosilicate glass spinner jars (Bellco Biotechnology)
Flat-plate stirrer platform (Bellco Biotechnology)

RCU IASI CENTRAL UNIVERSITY LIBRARY

Peristaltic pump (head size 15, 16, 24; with 0 to 600 rpm capability; Watson-Marlow or Cole-Parmer)
Sterile silicone tubing (Pt-cured; size 15, 16, 24; Masterflex from Cole-Parmer)
Sterile 50-ml syringes
Dissolved oxygen probe (Ingold)
pH probe (Ingold or Broadley James)

Additional reagents and equipment for growing mammalian cells in culture and counting viable cells by trypan blue exclusion (*APPENDIX 3C*), determining specific growth rate of cells (see Support Protocols 4, 5, and 9), and determining specific productivity of cells producing heterologous protein (see Support Protocols 6 and 7)

Establish cultures

1. Thaw a 1-ml cryovial of the culture in a 37 ± 1°C water bath for ~2 min with gentle shaking.
2. To a 250-ml screw-cap shaker flask, add a volume of cells from the vial sufficient to inoculate 100 ml of the medium to which the cells have been adapted (supplemented with L-glutamine and D-glucose and prewarmed to 37°C), at 1×10^5 viable cells/ml.

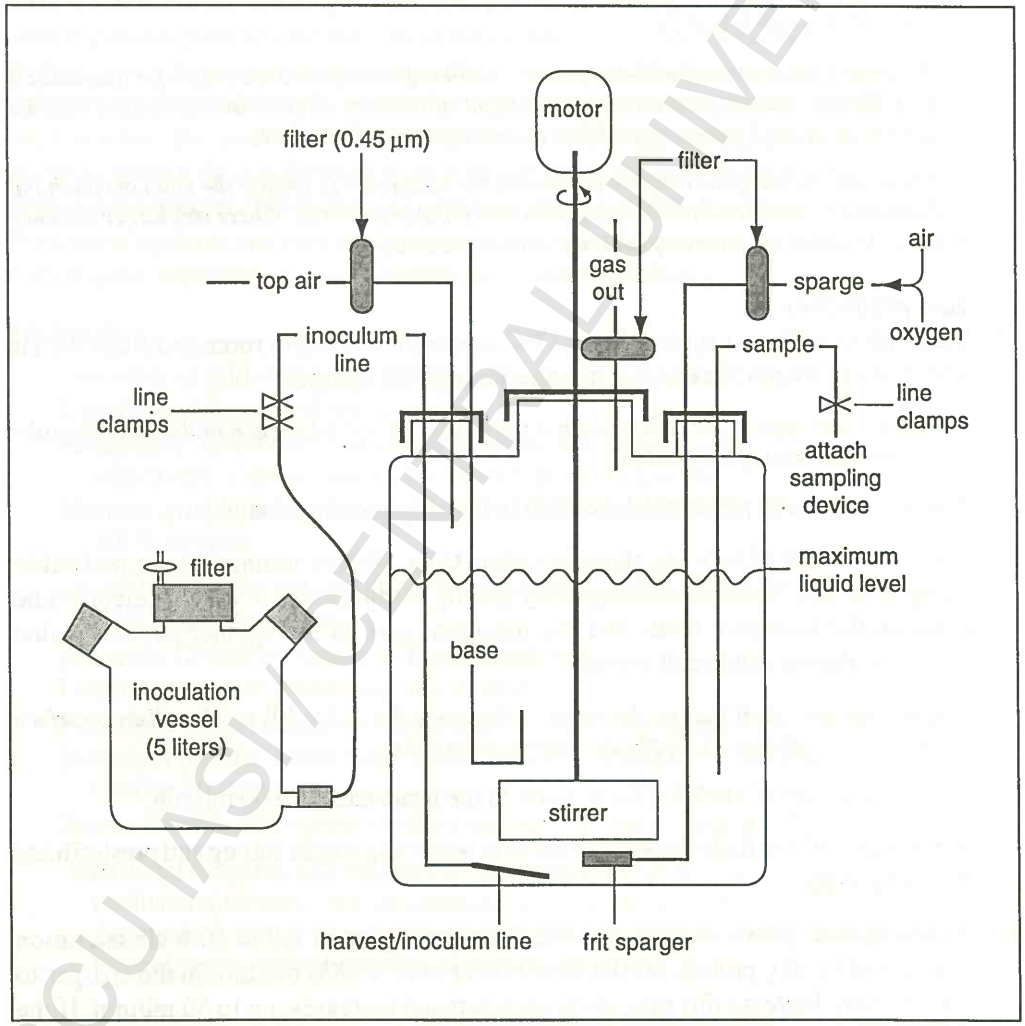


Figure 5.10.2 Typical arrangement of 20-liter spinner jar for batch cell culture.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Incubate flask in a humidified 37°C ± 0.5°C, 5% CO_2 incubator with rotation at 100 rpm on a platform shaker. Measure the viable cell density daily (*APPENDIX 3C*).

When the culture reaches mid–log phase (~3 to 4 days) the cells are suitable for inoculation of a 1-liter spinner jar containing 400 ml medium.

Inoculate spinner jar cultures

3. Add 400 ml of 37°C medium to a preincubated (5% CO_2, 37°C) 1-liter glass spinner jar. Add 100 ml of the mid–log phase shaker flask culture through the side arm. Determine the post-inoculation viable cell density, then place the spinner jar in the incubator on a flat-plate stirrer platform and set the stirrer speed to 60 to 80 rpm. Incubate, measuring the viable cell density daily.

   When the culture reaches mid–log phase (~3 to 4 days) the cells are suitable for inoculation of a 5-liter spinner jar containing 1600 ml of medium. The culture may be continued in the 400-ml spinner culture by adding glucose to a final concentration of 2 g/liter and glutamine to a final concentration of 2 mM on days 4, 6, and 8, along with daily pH monitoring and adjustment to 7.0 to 7.4 using 0.5 N NaOH.

4. Add 1600 ml of 37°C medium to a preincubated 5-liter glass spinner jar. Add the 400 ml of the mid–log phase culture. Determine the post-inoculation viable cell density and incubate with stirring as in step 3. Measure the viable cell density daily and monitor growth rate (see Support Protocols 4, 5, and 9) and specific productivity (see Support Protocols 6 and 7).

   When the culture reaches mid–log phase (~3 to 4 days) the cells are suitable for inoculation of a 20-liter spinner jar containing 4500 ml of medium. Otherwise the culture may be continued in the 5-liter jar by adding nutrients as described above.

   Knowledge of the growth curve will enable the researcher to choose the point at which the cells may be harvested with adequate viability and cell density. Where still larger volumes up to 16 liters are required, continue with step 5 onwards.

Increase production

5. Preincubate a 20-liter spinner jar at 37°C overnight in a warm room and flush the jar with filtered sterile 5% CO_2/air mixture through the sparger.

   The 20-liter spinner jar is fitted with a sterilizable dissolved oxygen probe and pH probe for measurement of dO_2 and pH.

6. Add 4.5 liter fresh prewarmed medium to the empty preincubated jar.

7. Transfer 1.5 liter of mid–log phase inoculum to the 20-liter spinner using a peristaltic pump with size 16 silicone tubing or by gravity feeding using a size 16 silicone tube between the inoculum bottle and the inoculum port on the spinner jar. Determine post-inoculation viable cell density.

   After inoculation, if the cell density is in the range 0.5–1.5 × 10^5 viable cells/ml, surface aeration is sufficient to supply the oxygen to the cells.

8. Set the flow rate of air/5% CO_2 mixture in the headspace to 100 ml/min.

9. Sample the culture daily through a side-arm port using sterile tubing and presterilized 50-ml syringes.

10. As the culture grows, and the dissolved oxygen begins to fall to 50% air saturation (measured by dO_2 probe), set the flow rate of air/5% CO_2 mixture in the sparger to ~25 ml/min. Increase this rate, as oxygen demand increases, up to 50 ml/min. If the demand for oxygen is not met (i.e., if the dO_2 remains <50% of saturation) using the air/5% CO_2 mixture, add 100% oxygen to the mixture. Maintain the dissolved oxygen

ECU IASI CENTRAL UNIVERSITY LIBRARY

between 20% and 100% air saturation. Maintain pH in the range 7.2 to 7.4 using automatic pH control or daily manual addition of 0.5 N NaOH.

11. When the culture reaches mid–log phase (around day 4) add 6 liters of fresh prewarmed (37°C) medium using a peristaltic pump and size 16 silicone tubing (see step 7).

12. When the 12-liter culture reaches mid–log phase (around day 6) add 4 liters of fresh prewarmed (37°C) medium.

The 16-liter culture may be grown up to the desired point for harvesting or may be operated in a feed-and-bleed mode. In feed-and-bleed mode, typically 4 to 10 liters of medium can be harvested from the late log-phase culture and replaced with the equivalent volume of fresh medium.

13. Harvest the protein product (see Basic Protocols 8 and 9).

GROWTH OF CELLS IN LARGE-SCALE BATCH REACTORS

BASIC PROTOCOL 7

Stirred-tank bioreactors may range in scale from 10 liters to 200 liters in a laboratory/pilot plant location. Bioreactors are distinguished from glass jar spinners and similar devices by their robust stainless-steel design and automatic control systems for temperature, dissolved oxygen, and pH, as well as instrumentation for monitoring important parameters such as medium flow rate, gas flow rate, pressure, weight, and liquid level. Bioreactors should possess ports around the base of the vessel for insertion of pH and dissolved oxygen probes. Since bioreactors may contain large volumes of culture and may operate at high cell density, the total oxygen demand is larger than can be supplied by agitation alone. In such reactors the addition of oxygen is achieved through sparging of the culture. This involves adding the gas through a frit-type device at the base of the reactor, or through a drilled-hole sparger. The authors recommend the frit design for reactors up to 500 liters. Bioreactor systems are supplied by specialized manufacturers of process equipment and may require lead time of 3 to 6 months from order to delivery.

Materials

Inoculum of mid–log phase cells in spinner jar (see Basic Protocol 6)
Ultrafiltered deionized water
Medium to which cells have been adapted (see Basic Protocol 5), supplemented with 8 mM L-glutamine and 4.5 to 5.5 g/liter D-glucose
Medical-grade, sterile-filtered 95% N_2/5% CO_2 and 95% air/5% CO_2 mixtures and 100% oxygen
10- to 100-liter bioreactor (New Brunswick Scientific, Bioengineering, Braun, New MBR, Applikon)
pH probe (Broadley James or Ingold), calibrated
Dissolved oxygen probe (Ingold), pretested
Autoclave large enough to accommodate bioreactor vessel (Finn Aqua)
Inoculum transfer bottle with bottom side port attached to sterile size 16 tubing (Bellco Biotechnology)
Sterile flexible disposable medium and harvest bags (Stedim Laboratories)
Additional reagents and equipment for counting viable cells by trypan blue exclusion (*APPENDIX 3C*), determining specific growth rate of cells (see Support Protocols 4, 5 and 9) and determining specific productivity of cells producing heterologous protein (see Support Protocols 6 and 7)

BOSTON UNIVERSITY CENTRAL LIBRARY

Prepare the bioreactor assembly

1. Assemble bioreactor components—e.g., agitator, sparger lines, outlet gas line, top headspace gas inlet, base addition lines, medium addition lines, inoculum transfer lines, harvest lines, and sample port lines.

   Recommended equipment includes sterile tubing (Pharmed) and connection tubing welder SCD IIB (Terumo Medical).

2. Prepare an inoculum for inoculation of the bioreactor.

   This would be between 0.1 and 0.15 liter of mid–log phase culture per liter of bioreactor operating volume. For a 100-liter bioreactor, an inoculum volume of 10 to 15 liters is required and would be produced by the 20-liter spinner method (see Basic Protocol 6).

3. Calibrate the pH probe and insert into reactor through the side port. Insert the pretested dissolved oxygen probe. Fill the vessel with ultrafiltered deionized (UFDI) water to cover the probes if sterilization is not immediately scheduled. Prior to sterilization, remove water.

4. Add ~10 ml ultrafiltered deionized water per liter of vessel volume and sterilize the vessel in an autoclave for 1 hr at 121°C, using slow exhaust for cooling phase.

5. Pump the culture medium into the vessel to cover the probes. Equilibrate system by sparging the medium and headspace at ~0.005 vessel liquid volumes per minute with a 95% N_2/5% CO_2 mixture, and allow the medium to reach the 37°C operation temperature. After equilibration (~5 hr required) is complete (as judged by an *x*-axis asymptotic dissolved oxygen trace), set the dissolved oxygen probe to 0% air saturation, then sparge the vessel with 95% air/5% CO_2 mixture. After equilibration, set the dissolved oxygen probe span to 95% air saturation.

   Optional recommended equipment includes a YSI 2700 Select glucose, glutamine, and lactate analyzer (YSI Instruments) and an IBI Biolyzer for ammonia and lactate (Kodak Lab Research Products).

Inoculate the bioreactor

6. Adjust the medium volume to 40% of the target final harvest volume (e.g., 4 liters in a 10-liter-operating-volume vessel). Equilibrate to 37°C, then inoculate to a starting density of 1 to 2×10^5 viable cells/ml using the inoculum transfer bottle.

7. Set the sparge air-flow rate to ~0.005 vessel volumes of air per vessel liquid volume and set the dissolved oxygen controller to add pure (100%) oxygen to the sparger on demand. Purge the headspace 0.0025 vessel liquid volumes of air. Agitate the culture at an rpm producing a tip speed of 1000 to 1500 cm/sec.

8. Continue incubation, determining the post-inoculation viable cell density daily (*APPENDIX 3C*). Monitor growth rate of cells

9. Feed medium beginning on day 2 and adjust the rate daily to give a rate of 2 liters per day per 1×10^9 total viable cells until the target volume is reached. At that point, stop feeding fresh medium.

10. During the initial-feed batch and final batch phases maintain the pH in the range 7.2 ± 0.1, the temperature at 37°C ± 0.5°C, and the dissolved oxygen in the range of 20% to 100% air saturation.

    The feed rates can be set through use of precalibrated pumps and checked using load cells that measure the medium supply container weights.

11. Harvest the protein product (see Basic Protocols 8 and 9) using a sterile flexible bag.

FCU IASI / CENTRAL UNIVERSITY LIBRARY

GROWTH OF CELLS IN CONTINUOUS CULTURE IN BIOREACTORS

Continuous culture refers to the operation of a bioreactor where the cells are removed at a particular rate and replaced with fresh medium at an equal rate. The ratio of the medium-addition rate to the culture volume is known as the dilution rate. Continuous culture may be desirable where large quantities of protein product are required for further studies, and is also of interest to the researcher for examining the potential effects of growth rate on the product. In this situation, a continuous reactor is set up with no cell retention devices. Where the researcher wishes to maximize the productivity of the given system, a continuous perfusion (cell retention) system would be used. Medium is removed at a rate corresponding to the addition rate of fresh medium but a large proportion of cells are retained in the reactor through use of suitable retention devices. The retention devices are usually either internal spin filters (supplied with bioreactors), external tangential cross-flow microporous filtration devices (Millipore, A/G Tech, Sartorius, Pall-Filtron) with cell recycle, external centrifugal devices (Centritech from Sorvall) with cell retention, or inclined settling surfaces (custom designed), which may be internal or external. Recently acoustic settlers have become available (Sonosep) and these are possible candidates for bioreactor scales around 10 to 30 liters.

The reactor assembly would incorporate the required peripherals for continuous operation (Fig. 5.10.3). In the case of perfusion, the settling device would be incorporated and sterilized along with the bioreactor. The basic setup is the same as described previously (see Basic Protocol 7).

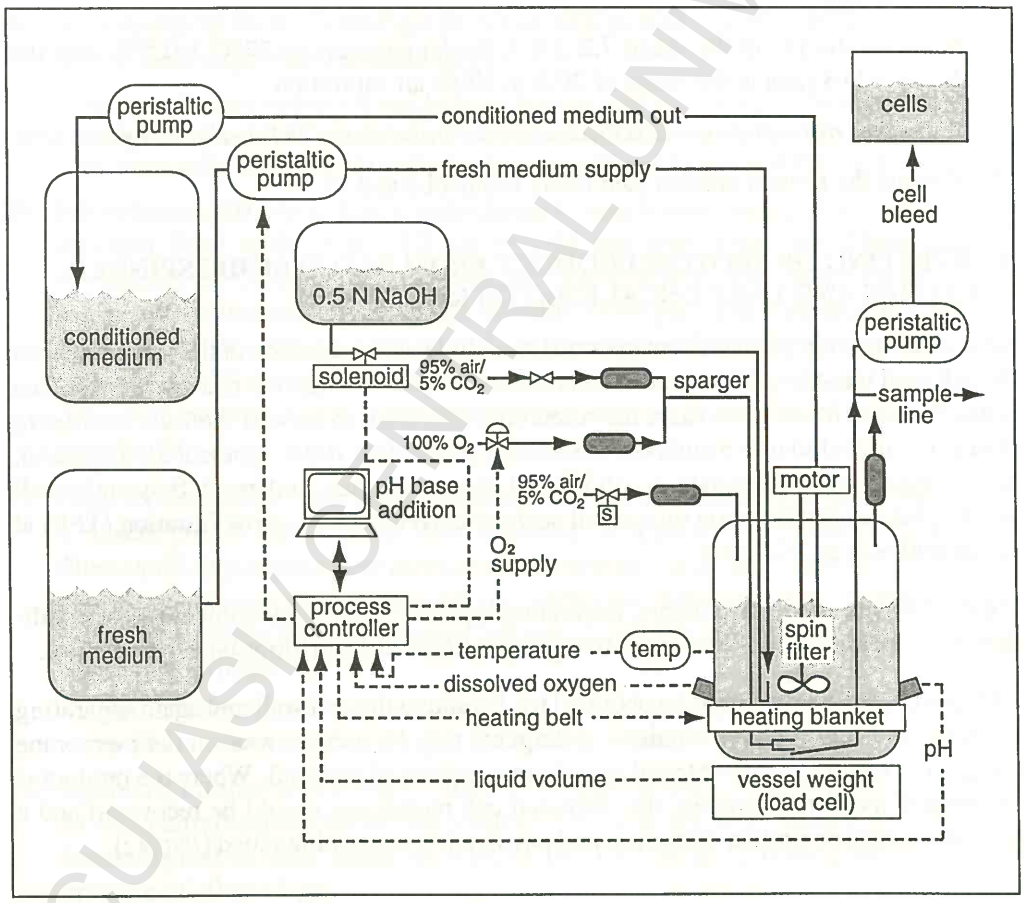


Figure 5.10.3 Compact-scale perfusion cell culture system. Shows typical connection of medium and product tanks to compact (100-liter) reactor. Simple pH, dissolved oxygen, and vessel weight control loops are illustrated. Medium and product vessels (for conditioned medium and/or cells) should be set up in a cold-box environment.

CU IASU CENTRAL UNIVERSITY LIBRARY

Additional Materials (also see Basic Protocol 7)

Bioreactor with accessories for continuous operation (consult New Brunswick Scientific, New MBR, Bioengineering AG, Braun, or other fermenter manufacturer; for perfusion operation where retention is required, use reactors with spin filter design, e.g., New MBR or Braun)

Sterile bags (Stedim Laboratories) or sterilized stainless steel containers for collection of medium

1. Set up and inoculate bioreactor (see Basic Protocol 7, steps 1 to 7). For the simple continuous mode and perfusion operation, feed medium at a rate of 2 liters per day per 1×10^9 total viable cells, up to the total operating volume of the reactor.

2. When the cells have reached mid to late log phase, begin cell removal at the target rate corresponding to the desired dilution rate (the volume of medium added to the vessel per day per reactor operating volume)

   The reactor operating volume is constant with respect to time. Fresh medium is fed to the reactor to maintain constant volume. In the case of the perfusion reactor, the retention device results in cells being returned to the bioreactor medium, and the cell density increases until it becomes limited by a nutrient or by toxic waste products such as ammonia and lactate. At that point a steady-state condition will be reached and cell density will remain constant.

3. Collect the conditioned medium as it removed from the bioreactor in sterile bags or sterilized stainless steel containers. Cool the stream to 4° to 8°C to minimize product degradation.

4. Maintain the pH in the range 7.2 ± 0.1, the temperature at 37°C ± 0.5°C, and the dissolved oxygen in the range of 20% to 100% air saturation.

   The feed rates can be set through use of precalibrated pumps and checked using load cells.

5. Harvest the protein product (see Basic Protocol 8 and 9).

HARVESTING OF PROTEIN PRODUCT FROM BATCH MODE SPINNER CULTURES AND LARGE-SCALE BATCH REACTORS

Most often, protein production in mammalian cells leads to secretion of the product across the cell wall into the culture medium where it accumulates with time in a batch mode (see Basic Protocol 8). In some cases the researcher may wish to harvest the cells in order to obtain an intracellular or membrane-associated protein (see Basic Protocol 9). In general, all strategies will necessitate a solid/liquid separation step and most frequently will employ batch centrifugation on a small scale or tangential cross-flow filtration (TFF) at larger scales, e.g., ≥20 liters.

For continuous-mode bioreactors, harvesting typically involves accumulation of conditioned medium, which is processed using depth filters and ultrfiltration concentration.

Cell-associated product must be obtained by disrupting the cell wall and again separating the solid and liquid phase. Nonionic detergents may be used to weaken the membrane and aid the release of non–integral-membrane–associated material. Where the product is an integral membrane protein, the disrupted cell membrane should be recovered and a typical membrane protein extraction and purification protocol pursued (*UNIT 4.2*).

PCU IAST / CENTRAL UNIVERSITY LIBRARY

Harvesting a Secreted Product

Materials

Cells in batch culture (see Basic Protocols 6 or 7 or Alternate Protocol 3)
Phosphate-buffered saline (PBS; *APPENDIX 2E*)
TE buffer or PBS (*APPENDIX 2E*)

Sterile bags (Stedim Laboratories)
Refrigerated centrifuge: Sorvall RC-5B or RC-3B, Beckman J2-21 or S600, or equivalent
Microfiltration system with membranes for cell concentration and debris/product separation (Sartorius, Pall-Filtron, or Millipore)
Depth filtration membranes for large-scale debris separation prior to sterile filtration (Cuno Bio-Cap 90 SP or equivalent products from Millipore or Sartorius)
Filter-sterilizing capsules (0.2-μm membrane; Sartorius, Pall, or Millipore)
Silicone tubing, autoclaved
Sterile polycarbonate containers
Ultrafiltration system with membranes for protein concentration (MWCO 10-kDa; Sartorius, Pall-Filtron, or Millipore)

1. Following completion of the batch process in spinner jar or a bioreactor remove medium from the reactor harvest port using sterile silicone tubing and a peristaltic pump. Collect into a sterile container such as a presterilized glass jar or sterile disposable plastic flexible bag. Chill collected material to 4° to 10°C

2a. *For volumes ≤6 liters:* Centrifuge 10 min at 1000 to 4000 × *g*, 4° to 10°C, in a Sorvall RC-3B centrifuge with standard rotors.

For up to 20-liter volumes it may be feasible to use a larger-capacity laboratory centrifuge such as the Sorvall RC-5B fitted with the swinging bucket rotor accommodating six 1-liter buckets; using this setup the cells should be centrifuged 15 min at 4000 × g, 4° to 10°C.

2b. *For volumes >20 liters:* Use a wide-channel nonscreened microporous cross-filtration unit fitted with ~1 ft^2 (0.2 μm or 0.45 μm pore rating) per 2-liter culture. Recirculate the cells through the retentate loop to give an inlet gauge pressure ≤10 psi. Limit the transmembrane gauge pressure to <4 psi using a peristaltic pump on the permeate line. Concentrate the cells in the retained filter stream by reducing the volume 10-fold, then permeate a further 3 vol PBS to wash the cells and recover the remaining product. Collect and chill the permeate in sterile plastic bags or sterile containers.

3. Filter the centrifuge supernatant, perfusion conditioned medium, or permeates using a rinsed, prewet liquid depth filter followed by a prewet or rinsed 0.2-μm rated filter-sterilizing capsule. Use autoclaved silicone tubing on the sterile side of the filter capsule and recover product in sterile polycarbonate containers or sterile plastic bags. Rinse the capsules with a minimum amount (50 to 100 ml) of TE buffer or PBS to recover the product.

IMPORTANT NOTE: *It is very important to prewet or rinse capsule filters as they hold significant amounts of liquid, which can result in a loss of product.*

4. *Optional:* Further concentrate filtrate by tangential ultrafiltration using a 10-KDa MWCO ultrafiltration membrane sized to 1 ft^2 of screened membrane per 20 liters. Recirculate at an inlet gauge pressure of 20 to 30 psi and a transmembrane gauge pressure of 10 to 15 psi.

A minimum recirculation loop volume of 100 ml is feasible for small-scale systems (Millipore, Sartorius, or Filtron Slice units) and will define the maximum physical concentration factor.

BOURKE UNIVERSITY LIBRARY

5. Rinse the filtration loop and filters with 50 to 100 ml of TE buffer or PBS to recover the product.

The product may be purified directly or stored frozen (see Support Protocol 3) until required.

BASIC PROTOCOL 9

Harvesting a Cell-Associated Product

Materials

Cells in batch culture (see Basic Protocols 6 or 7 or Alternate Protocol 3)
Phosphate-buffered saline (PBS; *APPENDIX 2E*)
Homogenization buffer (see recipe), 4°C
Cell homogenizer: e.g., Microfluidizer (Microfluidics), Tissumizer (Tekmar-Dohrmann), Dounce homogenizer (Bellco Biotechnology), or Manton-Gaulin-APV homogenizer (APV-Gaulin)

Additional reagents and equipment for centrifugation or cross-flow microfiltration of cell suspension (see Basic Protocol 8)

1. Subject cell suspension from batch culture to centrifugation or cross-flow microfiltration (see Basic Protocol 8, steps 2a or 2b), but retain centrifuge pellets or cross-flow filtration–concentrated slurry. Keep cells at 4° to 10°C.
2. Resuspend cells in 3 to 5 vol PBS (or other isotonic buffer) and centrifuge 10 min at 1000 to 4000 × *g*, 4° to 10°C.

   The cell pellet may be stored frozen (<–35°C) or processed directly.

3. Resuspend pellet in 2 to 3 vol of 4°C homogenization buffer. Lyse cells using a tissue homogenizer at 4°C.

   For larger volumes, it is possible to recirculate the suspension through a gear-type pump at high flow rate (obtain 10-20 psig pump outlet pressure in a 0.25-0.5 in.-diameter tube recirculation loop).

4a. *For volumes <20 liters:* Centrifuge the homogenate 30 min at 5000 × *g*, 4°C to remove the bulk of the debris. Filter sterilize the supernatant, which will contain the released soluble protein, prior to storage or further processing.

4b. *For large volumes:* Remove the debris using microfiltration/tangential-flow filtration (MF-TFF) at an inlet gauge pressure of 20 psi and limit the transmembrane gauge pressure to <4 psi using a peristaltic pump on the permeate line. Wash the retained concentrated debris with 3 to 5 vol buffer. Filter-sterilize the permeate, which will contain the released soluble protein, prior to storage or further processing.

Cell debris containing integral membrane protein may be recovered as the retentate concentrated slurry.

The permeate stream may be concentrated using cross-flow ultrafiltration, as described for secreted protein (see Basic Protocol 8).

PCU IASI / CENTRAL UNIVERSITY LIBRARY

FREEZING CELLS FOR PRODUCTION OF CELL BANKS

In order to be able to produce recombinant proteins consistently from batch to batch, it is advisable to produce a number of equivalent storage vials of the cell line—often referred to as a parent or master cell bank. Before attempting large-scale culture with a new cell bank the researcher is strongly advised to ensure that the cell bank is not contaminated by microbes, including mycoplasma, and that it grows normally (i.e., compared to the growth curve obtained during preparation). A growth curve test of the vial would detect contamination by bacteria, yeast, and fungi, while mycoplasma may be determined using commercial kits such as the Fluorescent DNA stain (Hoechst). A primary cell bank, or parent or master cell bank (MCB), can be used as the source to create subsequent working cell banks (MWCB) when necessary, depending on the scale of production. General principles and procedures for cryopreservation of attachment-dependent cells are provided in *APPENDIX 3C*. The following protocol is applicable to cells that have been adapted for suspension growth.

Materials

High-viability (>95%) log-phase culture of cells
Freezing medium: complete medium used for growing cells supplemented with 10% to 20% (v/v) FBS and 5% to 10% (v/v) DMSO (Sigma), 4°C

45-ml sterile screw-cap tubes polycarbonate centrifuge tubes (Falcon from Becton Dickinson)
Sorvall T-6000B centrifuge (DuPont) or equivalent
1.5-ml Nalgene cryotubes (Nunc)

Additional reagents and equipment for counting viable cells by trypan blue exclusion (*APPENDIX 3C*)

1. Take an aliquot of ~1 ml from a high-viability log-phase culture and determine the viable cell density (*APPENDIX 3C*). Determine the number of vials to be produced, then, based on the viable cell density and the total number of frozen vials to be prepared for the cell bank, calculate the volume of suspension culture needed.

   Each vial should usually contain 0.5 or 1×10^7 viable cells with a density of 1×10^7 cells/ml.

2. Transfer the desired volume of cell suspension to one or more sterile 45-ml polycarbonate centrifuge tubes and centrifuge 10 min at 300 to 350 × *g* in a T-6000B or equivalent centrifuge, room temperature.

3. Decant supernatant from the centrifuge tube(s) and resuspend pellet in required volume of 4°C freezing solution for a final concentration of 1×10^7 viable cells/ml.

4. Transfer 0.6- or 1.1-ml aliquots of the cell suspension into labeled cryotubes.

5. Cool and freeze cells at a rate of −1°C/min.

   Three general methods can provide this freezing profile: (1) the vials can be placed in a styrofoam box with a wall thickness of ~15 mm, which is then placed in a −70°C freezer; (2) the vials can be placed in a special cell-preservation container supplied by Nalgene which is filled with isopropanol and placed at −70°C; (3) an expensive programmable cooler can be used that can sense the temperature of the vial and feed in liquid N_2 accordingly.

   Placing the vials directly into liquid N_2 can result in the formation of ice crystals in the cells and poor viability when thawed.

6. When vials reach −70°C (or after 2 days) transfer to liquid N_2 for long-term storage.

BCU LRS / CENTRAL UNIVERSITY LIBRARY

SUPPORT PROTOCOL 4

NUMERICAL DETERMINATION OF SPECIFIC GROWTH RATE OF CELLS

Specific growth rate of cells any given time is defined as the rate of total biomass production divided by viable cell density at that given time:

$$\mu(t) = \frac{1}{V(t) \cdot X(t)} \left(\frac{d(V \cdot X)}{dt} \right)$$

where $\mu(t)$ is the specific growth rate at time t, $V(t)$ is the volume at time t, and $X(t)$ is the cell density at time t.

Specific growth rate is a function of cell type as well as the depletion of growth-limiting nutrients in the medium of choice and/or the accumulation of growth-inhibitory waste products in the medium. Both graphical (see Support Protocol 5) and numerical methods have been used to determine the specific growth rate of a culture. The growth rate is determined numerically by the following procedure.

1. Take a 1-ml sample from the culture and count viable cells (*APPENDIX 3C*) at time point 1.
2. After 4 to 24 hr (depending on how fast cells are growing) take another 1-ml sample from the culture to count viable cell density at time point 2.
3. Calculate the specific growth rate μ of cells as:

$$\mu_{\text{time 2}} = \frac{\ln\left(\frac{\text{viable cell density at time 2}}{\text{viable cell density at time 1}} \right)}{\text{time 2} - \text{time 1}}$$

SUPPORT PROTOCOL 5

GRAPHICAL DETERMINATION OF SPECIFIC GROWTH RATE OF CELLS

When complete growth curve data are obtained, the data may be input to a spreadsheet/graphics-type program capable of curve fitting and defining a polynomial best fit. The information can be used to determine the derived parameter of specific growth rate. For definition of specific growth rate, see Support Protocol 4. The growth curve is determined graphically as follows.

Materials

Spreadsheet software with graphical output (e.g., Microsoft Excel or Lotus 123)

1. Take a 1-ml sample from the culture for viable cell count (*APPENDIX 3C*) every 4 to 24 hr, depending on how fast the cells are growing, until late log or stationary phase.
2. Plot the viable cell density versus time on an *x-y* chart using spreadsheet software capable of graphical output.
3. Curve-fit the chart to a best 3- or 4-order polynomial function.

   The logarithmic growth phase (days 0 to 5, typically) can also be fitted with a log function or linear fit on a log scale.

4. Calculate the specific cell growth rate at time = T hr as the first-order derivative of the function divided by the viable cell density at time = T hr.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

NUMERICAL DETERMINATION OF SPECIFIC PRODUCTIVITY OF CELLS PRODUCING HETEROLOGOUS PROTEIN

SUPPORT PROTOCOL 6

The specific productivity of a protein at any given time is defined as the rate of increase in the protein mass divided by total viable cell mass at that given time:

$$\text{specific productivity } (t) = \frac{1}{X(t)}\left(\frac{dP(t)}{dt}\right)$$

where $X(t)$ is the cell concentration at time t and $P(t)$ is the product concentration at time t. Specific productivity is calculated as follows.

1. Take a 1-ml sample from the culture, count viable cells, and determine the product protein concentration by ELISA at time point 1.
2. After 4 to 24 hr (depending on how fast cells are growing), take another 1-ml sample from the culture, count viable cells, and determine product protein concentration by ELISA at time point 2.
3. Calculate the specific productivity (q_p) by cells as:

$$q_p(\text{time } 1 \rightarrow \text{time } 2) = \frac{\text{protein concentration at time } 1 - \text{protein concentration at time } 2}{(\text{time } 2 - \text{time } 1) \times X_{avg}(\text{time } 1 \rightarrow \text{time } 2)}$$

where

$$X_{avg}(\text{time } 1 \rightarrow \text{time } 2) = \frac{\text{viable cell density}_{\text{time } 2} - \text{viable cell density}_{\text{time } 1}}{\ln(\text{viable cell density}_{\text{time } 2} - \text{viable cell density}_{\text{time } 1})}$$

GRAPHICAL DETERMINATION OF SPECIFIC PRODUCTIVITY OF CELLS PRODUCING HETEROLOGOUS PROTEIN

SUPPORT PROTOCOL 7

For definition of specific productivity, see Support Protocol 6.

Materials

Spreadsheet software with graphical output (e.g., Microsoft Excel or Lotus 123)

1. Sample cultures at time points 1 and 2, count viable cells, and determine protein concentrations (see Support Protocol 6, steps 1 and 2).
2. Plot the viable cell densities versus time and plot the protein concentrations versus time on *x-y* formatted charts using spreadsheet software capable of graphical output.
3. Curve-fit the charts with the best 3- or 4-order polynomial functions.
4. Calculate the specific productivity as the first-order derivative of the protein concentration curve-fit function divided by the viable cell density at any given time point.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

SUPPORT PROTOCOL 8

PREPARATION OF SAMPLES FOR PRODUCT ASSAY

Samples taken from the culture need special preparation for protein assay and storage. The handling procedures are different for cultures that make intracellular product and those that make secreted product. Freezing whole-cell cultures and subsequent thawing may affect the value and consistency of the assay in some cases. If samples must be stored frozen as whole-cell cultures prior to assay, the effects of freeze/thawing should be evaluated, but this is not recommended as a consistent method. In the procedures below, the culture is centrifuged or filtered to separate the cells (which are lysed to produce a cell extract if the product is intracellular) from the supernatant (which is handled separately when the product is secreted).

Materials

Sample of cell culture (see Basic Protocols 6 or 7 or Alternate Protocol 3)
0.9% NaCl, filter-sterilized, 4°C

Refrigerated centrifuge
Ultrasonic homogenizer (Misonix)

1. Centrifuge 1 to 5 ml of sample 10 min at 300 to 350 × *g*, 4°C. Decant supernatant; retain for assay or store frozen at ≤30°C for future analysis if protein product is secreted. If protein product is intracellular, proceed with step 2 and subsequent steps.
2. Resuspend the cell pellet in 1 ml of cold 0.9% NaCl solution.
3. Centrifuge again for 10 min at 300 to 350 × *g*, 4°C.
4. Decant supernatant and resuspend pellet in 1 ml of 4°C 0.9% NaCl.
5. Keeping the sample tube in an ice bucket, place the sample tube under the ultrasonic homogenizer and immerse the microtip probe in the cell suspension. Operate the sonicator for a 10-sec cycle with continuous power-on time of 0.5 sec and continuous power-off time of 1.5 sec, with the power set at low to medium level.

   After the completion of a 10-sec cycle, load a drop of cell suspension onto a hemacytometer slide and observe the lysate under 125× magnification using phase contrast. If insufficient disruption is observed, perform further 10-sec sonication cycles until cells lysed.

6. Centrifuge the suspension of disrupted cells 10 min at 800 to 1000 × *g*, 4°C.
7. Retain supernatant for assay or store frozen at ≤30°C for future analysis.

SUPPORT PROTOCOL 9

DETERMINATION OF SHAKER FLASK GROWTH CURVE

A complete 10-day growth curve is useful for determining the expected cell yield in batch culture, for optimization of time of harvest, and for information making it possible to compare culture performance at different scales.

Additional Materials (also see Basic Protocol 6)

Thawed vial of cryopreserved cells (0.5 ml to 1 ml at 10^7 viable cells/ml) or aliquot of healthy cells in exponential growth phase (10^7 cells total in ~25 ml)

Additional reagents and equipment for determination of specific growth rate (see Support Protocols 4 and 5) and specific protein productivity of cells (see Support Protocols 6 and 7)

1. Preincubate a 250-ml shaker flask, containing sufficient medium such that the volume after adding the aliquot of cells will be 50 or 100 ml total, in a humidified 37°C, 5% CO_2 incubator. Add an aliquot of cells that will give a starting density of 0.5–1.0 × 10^5 viable cells/ml.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

2. Remove a 1.0-ml sample to determine post-inoculation viable cell density (*APPENDIX 3C*) and protein product titer (in supernatant or cells). Place the flask containing the rest of the cells on a rotating platform in the incubator and incubate with agitation at 80 to 100 rpm.
3. Sample 1.0 ml from the flask daily and determine viable cell density and protein product titer.
4. Feed cells by adding glutamine to a final concentration of 2 mM and glucose to a final concentration of 2 mg/ml on days 4, 6, and 8. Check the pH on a daily basis from day 4 onwards, and correct to pH 7.2 ± 0.1 if necessary using 0.5 N NaOH.
5. Keep a tally of sample volumes removed and the glucose, glutamine, and NaOH volumes added to enable calculation of the cell density dilution that has occurred as a result of additions and sampling.
6. Determine the specific growth rate (see Support Protocols 4 and 5) and productivity (see Support Protocols 6 and 7).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

αMEM medium

Purchase Minimal Essential Medium Alpha (αMEM; Irvine Scientific, Life Technologies, Sigma, JRH Biosciences, or HyClone) without L-glutamine and supplemented with 2.2 g/liter sodium bicarbonate. For growth medium, purchase formulation containing ribonucleosides and deoxyribonucleosides; for DHFR selection medium, purchase nucleoside-free formulation. Supplement with Pluronic F-68 (BASF) to a final concentration of 0.05% (w/v). Add 10 ml of 200 mM L-glutamine stock solution (HyClone; store at –20°C) per liter of medium immediately prior to use. Supplement as required by protocol with γ-irradiated fetal bovine serum (FBS; HyClone; store at –20°C prior to addition), heat-inactivated 30 min at 56°C. For DHFR selection, use dialyzed FBS (HyClone) to ensure absence of nucleosides. Store prepared medium (without the serum) up to 1 month at 4° to 8°C.

DMEM/F12 custom-modified medium

DMEM/F12 medium (1:1 DMEM/F12; e.g., Life Technologies, JRH Biosciences, HyClone, or Irvine Scientific), without thymidine, hypoxanthine, or L-glutamine, containing the following (final concentrations):
25 mM HEPES
5 g/liter D-glucose
0.11665 g/liter $CaCl_2$ (anhydrous)
0.04884 g/liter $MgSO_4$ (anhydrous)
0.061 g/liter $MgCl_2 \cdot 6H_2O$
2.5 g/liter sodium bicarbonate
0.05% (w/v) Pluronic F-68 (BASF)
Osmolality 280 to 320 mOsm/kg, adjusted with NaCl
Store up to 1 month at 4° to 8°C
Add L-glutamine to 4 mM final concentration just prior to use
Supplement as required by protocol with γ-irradiated, dialyzed fetal bovine serum (FBS), heat-inactivated 30 min at 56°C.

This formulation (minus the L-glutamine and serum) may be prepared to order by Life Technologies, JRH Biosciences, HyClone, or Irvine Scientific.

BOSTON CENTRAL UNIVERSITY LIBRARY

Homogenization buffer

Appropriate buffer (e.g., 50 mM Tris·Cl, pH 7.5; *APPENDIX 2E*) containing:
1 mM EDTA
0.1 to 1.0% Tween 80 (omit for extraction of soluble intracellular proteins)
Store up to 2 weeks at 4° to 8°C

Resuspension buffer

50 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*)
10 mM EDTA
100 µg/ml RNase A
Store up to 1 week at 4° to 8°C

COMMENTARY

Background Information

Mechanisms of expression

The level of expression of genes introduced into mammalian cells will depend on the gene copy number, the site of chromosomal integration, the transcriptional control elements, and mRNA stability. Availability of translational components—e.g., aminoacyl-tRNA, ATP, GTP, and other nutrients—is not likely to be limiting in mammalian cells growing under cell culture conditions.

Expression vectors are modular in design, with various elements such as promoter-enhancer sequences, transcription termination/poly(A) sequences, multiple cloning sites, selection marker sequences with promoters and poly(A) sequences, and bacterial elements for replication and maintenance. Constitutive expression systems are most frequently used for stable expression. These include the strong promoters such as the cytomegalovirus (CMV) promoter, the β actin promoter, the SR α promoter, and the Rous Sarcoma Virus (RSV) promoter, which are capable of high levels of transcription. Following transcription in the nucleus, the primary transcript is capped and processed to remove any introns and the poly(A) tail is added before the mRNA is transported to the cytoplasm. Since most recombinant proteins are isolated as cDNAs, they do not contain introns. The poly(A) signal is required at the 3′ end of the foreign protein gene sequence for the transport and stability of the mature transcript. Commonly used poly(A) sequences are BGH, SV40, and HSV-tk; these are found in commercially available vectors. Although many commercially available plasmids contain introns, these offer no significant advantage in tissue culture cells (Brinster et al., 1988).

Translation of the message in the cytoplasm involves scanning of the message by the 30S ribosome complex from the cap in the 5′ to 3′ direction until it reaches an AUG codon in a favorable context, which then initiates translation. Thus, alteration of the original context of translation initiation to the Kozak's consensus sequence—GCCA/GCCAUGG—in the DNA molecule, may improve expression through increased efficiency of translation initiation (Kozak, 1986). In the consensus sequence, the –3 position and the +4 position are most critical. Another essential requirement is a stop codon at the end of the foreign-gene cDNA for termination of translation. For efficient translation, the DNA sequence should be altered to incorporate preferred codons of the expression system of choice where there is great disparity in codon usage. The cDNAs should also contain the appropriate signal sequences in frame, either on the 5′ or 3′ end, depending on the targeted organelle (see Aitken, 1997).

The selection marker is the other most crucial part of the vector apart from the *cis* elements that regulate expression of the protein. The *cis* elements required in the transcription unit of the selection marker are similar to those used for the protein. When the goal is to obtain high-level expression, it is important for the selection to be stringent and for the expression of the gene of interest to be linked to the expression of the marker gene. Methods for obtaining stringent selection involve weakening the expression of the marker protein through altering the context of translation initiation (Kozak, 1986), using a weak promoter such as SV40, or decreasing the mRNA stability (Miloux and Lupker, 1994). Mutation of the protein sequence to decrease activity has been used as a means to increase the stringency of selection, for example in the neomycin phosphotransferase gene (Yenofsky et al., 1990). In

RCU IASRI CENTRAL UNIVERSITY LIBRARY

many earlier transfections, the marker gene and the protein gene were on separate plasmids. Current vectors carry both genes on the same plasmid, which provides better probability of cointegration in stable transfection. The single-vector strategy does not essentially guarantee that the marker and protein expression are correlated in cases of very low expression of the marker gene (Kaufman et al., 1985). This motivated the incorporation of internal ribosomal entry site (IRES) sequences from encephalomyocarditis (EMCV) virus (Kaufman et al., 1991) to allow linking of expression at the RNA level. Although the fusion-protein approach (*UNITS 5.1, 6.6 & 6.7*) provides the best linking of genes it is seldom applied to produce proteins for structural or biological characterization since it will pose issues such as the effect of the fusion on the structure and function of the fusion protein. Linking reduces the chance of vector rearrangement, which could result in nonexpression of protein of interest with continued expression of surrogate marker. The use of multiple elements that can be selected or screened for, such as neomycin and GFP together or neomycin and DHFR together, reduces the chance that clones with expression of two surrogate markers will end up with no production of the protein of interest.

Thus far the methods discussed assume random integration of the heterologous DNA into the genome. The random-integration approach results in a very low frequency of integration into a transcriptionally active locus. This frequency can be increased through targeted integration using retroviral flanking sequences (Wurm et al., 1996), or by utilizing DNA sequences that flank the gene of interest in high-producer cell lines (Fukushige and Sauer, 1992; Morris et al., 1997).

Selection and screening of transfected cells

Selection and screening are the two different methods that help isolate cells expressing the marker gene. The marker gene is an indirect indication that the protein of interest is also being expressed. In selection, cells that do not express the marker are killed by the drug, and in screening, the cells that express the marker are isolated through flow cytometric or other techniques. Marker genes for selection are most often drug-resistance genes or easily detectable protein markers such as green fluorescent protein (GFP). In addition, a membrane-targeted protein that may be detected at the cell surface by a fluorescent-labeled specific antibody may act as a useful marker. Examples of drug-resistance markers include hygromycin phosphotransferase, which confers resistance to hygromycin; neomycin phosphotransferase, which confers resistance to geneticin (G418); thymidine kinase, which confers resistance to aminopterin; xanthine-guanine phosphoribosyltransferase, which confers resistance to mycophenolic acid; adenosine deaminase, which confers resistance to 9-β-D-xylofuranosyl adenine; and a mutant dihydrofolate reductase, which confers resistance to high levels of methotrexate. Although several marker systems are available (Mortensen et al., 1997; Kriegler, 1990; Kaufman, 1990), the neomycin phosphotransferase system is most commonly used and is available in a number of commercially available vectors (Table 5.10.1).

Amplification provides the means to go beyond the selection stage and enables higher gene dosage to be obtained. This usually leads to increased specific production of the gene of interest. Amplifiable elements are provided by recessive markers, which may be transferred into host cells lacking the related cellular function. Common examples are genes for dihydrofolate reductase (*DHFR*), which allows the transfected cells to grow in methotrexate (MTX), and for glutamine synthetase (*GS*), which enables growth in the presence of methionine sulfoximine (MSX). It is interesting to note that these drugs inhibit activity of marker enzyme through tight stoichiometric binding, whereas most drugs used in selection are the substrate for the marker enzyme and are inactivated by it. Other less frequently used recessive markers are adenosine deaminase (*ADA*), asparagine synthetase (*AS*), and ornithine decarboxylase (*ODC*; Mortensen et al., 1997; Kaufman, 1990).

In the case of the recessive marker gene *DHFR* in CHO (*dhfr*$^-$) cells, the methotrexate concentration is increased (when cells are growing adequately) by stepwise increases up to 10μM. After each step increase in the methotrexate concentration, the cells adapt and grow through increased *dhfr* gene expression. The higher-producing clones are screened and sorted for further rounds of amplification. The *GS* recessive marker has been used in CHO cells, and high titers (>100 mg/L) of foreign proteins have been obtained after two rounds of amplification using MSX (Cockett et al., 1990). *DHFR* can serve as a selection/amplification marker in *DHFR*$^+$ cells, such as CHO-K1, through the use of strong promoters to drive its expression, or with the low MTX affinity–mutated mouse *dhfr* gene (Simonsen and

BOSTON CENTRAL UNIVERSITY LIBRARY

Levinson, 1983), or multiple selection markers (e.g., *neo*^{R}; Kaufman, 1990). In these instances, selection for stable integration of *dhfr* requires moderately high (200 nM) MTX level to overwhelm the endogeneous DHFR and therefore the amplification potential is limited.

Amplification provides a means to enhance the specific productivity of a cell, particularly where the level of expression is gene-dosage limited. Limits to the potential increased productivity exist and may be due to saturation of the translational capacity of the cells (Pendse et al., 1992) or to instability of gene copy number (Wurm, 1992). Thus, there is likely a point where the level of message is not limiting the overall expression of protein. Although high synthesis rates have been described, the overall expression rate may be limited by the secretory capacity of the cell (Cockett et al., 1990). Coexpression of a protein with higher levels of endoplasmic reticulum resident proteins (BiP, PDI, and calnexin) has been performed as a means to facilitate higher secretion rates, but there has been no general approach that guarantees higher secretion rate. Where cells are amplified to high gene copy number there is no absolute assurance that the amplicons are stable in the absence of the MTX. Following amplification, it is useful to subpassage the cells without MTX and determine the stability of production by detection of the protein of interest over several passages. Another method for evaluating the heterogeneity of gene copy number and stability of the foreign genes in the cell population is the use of fluorochrome-labeled MTX (MTX-F) and flow cytometry (Kaufman, 1978).

Although there will always be significant time investment in amplification of expression, time savings may be possible by accelerating the selection and cloning procedure by use of flow cytometric sorting of producer cells. Use of these methods requires marker that can be noninvasively detected through fluorescence-based methods. A unique reporter that satisfies these criteria is the green fluorescent protein (GFP) from jellyfish (Chalfie et al., 1994; Kahana and Silver, 1996). Cells that express this protein can easily be detected and sorted using flow cytometric techniques (Subramanian and Srienc, 1996). GFP can be substituted with membrane-targeted proteins that are also amenable to flow cytometric detection and sorting. For example, the pHOOK system, commercially available from Invitrogen (pHOOK-2 vector; Capture-Tec 2 kit) can be used (Mortensen et al., 1997; Chesnut et al., 1996). In this system, a single-chain antibody (sFv) against a hapten is expressed in addition to the transcription unit for the gene of interest. There are also two immunofluorescence tags, *c-myc* and *HA*, fused to the sFv. Antibodies against the sFv or the tags could be stained through direct immunofluorescence staining (primary antibody is conjugated with a fluorescent label, i.e., FITC or PE) or through indirect immunofluorescence staining (the primary antibody has no label but a secondary antibody that recognizes the primary is labeled). Alternatively, cells expressing the sFv can be separated using magnetic affinity cell sorting (MACS) where micron-size paramagnetic beads coated with the hapten are mixed with transfected cells and the cells to which the beads are bound are separated using a magnet. One disadvantage with the MACS is that it can only separate positive from negative but cannot differentiate between cells with various amount of surface proteins (density). This situation may eventually not be limiting, with the introduction of continuous magnetic separation devices that can fractionate on the basis of surface-marker density (Chalmers et al., 1998).

The use of surrogate markers can be completely avoided if the protein of interest can be directly quantitated and high-producer clones can be isolated. Such an assay, which measures product secreted from single cells encapsulated in agarose beads using an immobilized matrix for the capture and detection of the protein of interest, has been applied to the isolation of hybridomas with high antibody productivity (Gray et al., 1995).

Mechanisms of secretion

Secretion of protein from the cell into the medium is the most commonly used expression approach. In the case of glycoproteins, the routing of the protein through the cell is dependent on a leader sequence in front of the protein sequence for endoplasmic reticulum docking and uptake, and lack of the C-terminal endoplasmic reticulum–retention sequences. The protein will be transported from the endoplasmic reticulum (with addition of a lipid-linked oligosaccharide dolichol phosphate precursor and trimming of mannose and glucose) through to the Golgi (with trimming of mannose and addition of N-acetyl glucosamine, glucose, galactose, fucose, and sialic acid residues) and to the cell surface for secretion. In the absence of a leader sequence, the protein will remain in the cytoplasm where other motifs and chaperones may channel the protein to the nucleus, to the mitochondria, or into membranes.

It may also simply accumulate in the cytoplasm as unglycosylated protein. For more information on the various protein motifs, see Aitken (1997) and *UNIT 17.1*. Proteins that are channeled through the ER and Golgi of the cell are presented to the cell's glycosylation machinery. The compartmentalization and translocation mechanisms are briefly reviewed in *UNIT 5.9*.

Choice of cell lines

Chinese hamster ovary cells have become a widely used host for stable expression of heterologous genes. The CHO cell lines DG44 and DXB11 are *DHFR*$^-$ and commonly used to enable methotrexate selection and amplification of chromosomal integrated genes (see *UNIT 5.9*). Both these cell lines may be obtained from Dr. Lawrence Chasin of Columbia University (e-mail *lac2@columbia.edu*). Selection approaches, other than methotrexate, use CHO-K1 cells (ATCC #CCL-61).

A number of CHO cell mutants exist that enable the production of glycoforms with removal or addition of linkages or sugar types. The mutants were isolated by their ability to grow in the presence of specific toxic lectins. The resistant cells do not bind the toxic lectins on their cell surface. This is a consequence of the mutant cell line lacking a specific glycotransferase enzyme and the resulting lack of the related lectin-binding carbohydrate structure on the cell surface. A number of these mutant CHO cells have been isolated and described and can be grown in culture (Stanley and Ioffe, 1995), providing interesting potential options for tailoring the carbohydrate structures. A mutant designated as Lec3.2.8.1 can provide glycoproteins with all the N-linked carbohydrate as Man_{5-7} and the O-linked truncated to a single GalNAc (Stanley, 1989). Other approaches to in vitro modulation of the glycoprotein carbohydrate structures involve use of specific inhibitors such as tunicamycin, castanospermine, deoxynojirimycin, deoxymannojirimicin, and swainsonine, which inhibit specific steps in the endoplasmic reticulum such as the addition of the lipid-oligosaccharide dolichol phosphate precursor to asparagine, or the trimming of glucose and mannose residues (*UNIT 12.3*). If the inhibitors are toxic to cells at their functional level, then gradual adaptation of the cells to the compounds may be required.

Analysis of the detailed carbohydrate sequence and distribution of structures for purified glycoproteins is still challenging, although the advent of capillary zone electrophoresis and electrospray mass spectrometry have been useful for fractionation and analysis of glycoforms present in recombinant proteins (Parekh and Patel, 1992; Yim, 1991). Although there are no preparative-scale separation methods to resolve larger quantities of glycoforms to homogeneity, immobilized lectin columns are available that enable purification of high-mannose glycoprotein from heterogeneous mixtures (*UNIT 9.1*). Use of purified specific exoglycosidases may be useful for sequential and specific digestion of the carbohydrate structures to a desired species (Brady et al., 1994).

Choice of medium

CHO cells readily adapt to suspension growth in medium containing lower Ca^{2+} and Mg^{2+}, in the range of 0.2 to 0.6 mM respectively. Spinner MEM medium, with reduced calcium and increased phosphate, was first used for anchorage-independent growth in 1959. A commonly used medium is DMEM/F12 (Life Technologies, JRH Biosciences, or HyClone) supplemented with fetal calf serum (Life Technologies, HyClone). Following adaptation to suspension, cells may be subjected to stepwise changes in medium components—e.g., decreasing levels of serum. Adapting the cells to growth in medium containing low serum will generally make the cells grow with less clumping. Another benefit of growing in lowered serum is reduction of contaminant protein level, which makes the subsequent purification easier. Serum contains bovine transferrin and mitogenic growth factors. When the concentration of serum is <1% (v/v), addition of insulin or IGF to the medium is needed.

Mechanisms of cell growth

Cells may be grown in shaker flasks, spinners, or larger bioreactors. The choice of approach is dictated by the quantity of final purified product required (Table 5.9.3 and Table 5.9.4).

Bioreactors may be set up in glass spinner vessels fitted for pH and dO_2 control (Griffiths, 1992). Larger-scale operation requires use of stainless steel vessels, which are available from 10 to 10,000 liters. Scale-up to 100 liters is adequate for intermediate-scale production. Such reactors may be operated in batch, fed batch, or continuous mode. Batch operation is the least complicated approach and the one that is least vulnerable to contamination, since fewer manipulations are required. Highest titer levels are achievable in fed batch because of the high cell density and high specific growth rate. Perfusion reactors enable high product titer and

COLUMBIA UNIVERSITY LIBRARY

productivity, which result from the higher cell density achievable, although the lower target growth rate may decrease the specific productivity. Cell-specific productivity is usually higher in fast-growing cells because of an increase in the frequency of S-phase cells in culture, and the use of S-phase-active promoters (Gu et al., 1993). High density is obtained by operation in continuous mode with cell retention in the bioreactor. Cell-retention devices include internal or external inclined settlers, spin filters, or external microporous filtration devices. Recently, an acoustic cell settler (Sonosep) and a cell-recycling external centrifuge were developed (Centritech AB, DuPont) and have been made available for compact operation up to 40 liters.

Cell-retention devices should provide a cell retention efficiency of >75% to support a reactor steady-state cell density at a minimum specific growth rate of 0.3 per day. A target growth rate should be stipulated in continuous-mode strategies. Perfusion reactors require an initial fed-batch phase to establish high cell density. As the specific growth rate declines, the removal rate of cells is set to the target value (typically 0.25 to 0.4 per day).

Following completion of batch growth, or during continuous-mode operation of cell culture bioreactors, it is necessary to recover the product. The primary recovery steps may be designed to harvest the cells when the product is intracellular, or most commonly to harvest the conditioned medium containing the secreted product. The types of methods used are not very different from those typically used in bacterial or yeast bioprocessing. A summary of typical recovery operations used at small and large scale is given in Figure 5.10.4.

Mammalian cells are ~10 to 20 times larger in dimension than bacteria and possess a lipid bilayer cell wall, which renders the mammalian cell much weaker than the more complex cell walls of bacteria and yeast. The larger size provides for a sedimentation rate of 10 to 20 cm/min at 1 g, in the order of 100 times larger than that of bacteria; hence, low-speed centrifugation is able to separate the cells without compression and cell damage. Thus, recovery of the product from the culture may be readily accomplished at laboratory and compact bioreactor scale. Batch centrifuges capable of spinning at 2000 to 5000 × *g* can separate cells from the conditioned medium. For larger volumes (>20 liters), tangential-flow microporous membranes (0.2 μm pore size) may be used to obtain a permeate containing extracellular protein and a retentate containing the concentrated cells. It is important to minimize cell damage and product contamination and loss, which could result from exposure to shear forces or bursting bubbles. Cross-filtration devices are available for mammalian cell operations and feature open-channel design to minimize shear damage. These operate at low transmembrane gauge pressure (<5 psi) to avoid cell deformation and fouling at the membrane surface. Since most mid-scale-up to pilot-scale continuous centrifuges contain acceleration zones which can damage cells, they are rarely used in primary recovery of cells or conditioned medium. In dealing with large volumes of conditioned medium from perfusion reactors, it may be possible to allow the cells in the conditioned medium to settle overnight. The top 80% of the liquid could then be clarified by tangential flow filtration (TFF) or by depth-filtration membranes (Seitz, Cuno, FPI) and then concentrated using tangential cross-flow ultrafiltration.

If intracellular product is desired, it is necessary to disrupt the cell membrane to release the product. Typically, the first step in such a scheme would be to separate the cells from the culture medium to eliminate the serum contaminants. This is achieved by centrifugation or gravity separation of the cells overnight in a cold room, followed by decanting of the clear top liquid phase. The cells can be resuspended in a simple buffer containing a low concentration of a suitable nonionic detergent to weaken the cell membrane. Exposure of the cells to high laminar shear using a homogenizer (APV-Gaulin), or recirculation at high rates using a rotary displacement pump, will rupture the cells further and release the product. After cell disruption, the difficult step of separating the cell debris from the released product must be accomplished. For this step, high-speed centrifugation could be applied or TFF could be used at low transmembrane pressure and high recirculation rate employing screened membranes.

Throughout the recovery operations the product may be exposed to physical conditions, or degradative enzymes released by cell lysis, which may modify the protein. Typically, serine proteases are active at neutral pH and lysosomal-based proteins are active at low pH (4 to 6). Resistance to proteolysis may result from inherent structural properties of the proteins such as buried protease sites or the conditions of processing such as inclusion of protease inhibitors or speedy processing at <10°C. The presence of sialidase, as a result of cell lysis,

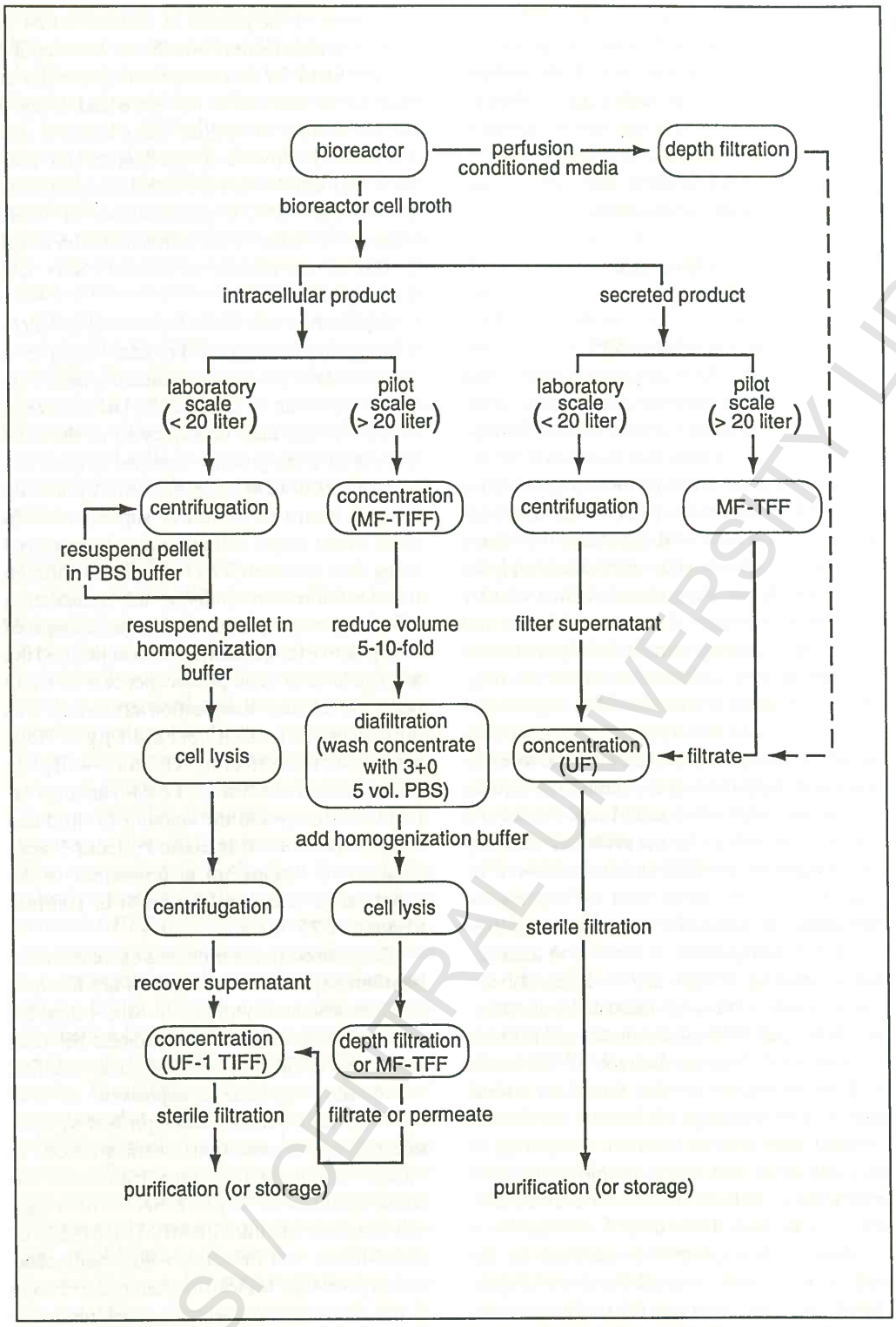


Figure 5.10.4 Cell culture product recovery flow schemes for laboratory and pilot scale processes. Options for intracellular and secreted proteins are shown. For details, see Basic Protocol 8. Abbreviations: MF, microfiltration; TFF, tangential-flow filtration; UF, ultrafiltration.

BCU IUSI / CENTRAL UNIVERSITY LIBRARY

may also modify the glycoprotein and decrease the sialic acid content. Processing quickly and keeping the product at 4° to 10°C is the basis for a recovery operation with minimal physical product damage. Where specific problems of aggregation, proteolysis, and degradation are encountered, the researcher may have to consider use of chemical inhibitors.

Critical Parameters and Troubleshooting

Expression vector construction

No molecular biology protocols have been provided for constructing an expression vector, since such information is widely available (e.g., Ausubel et al., 1998, and Sambrook et al., 1989). However, the following guidelines should be considered no matter which approach is used. Construction of the vector with most commercial vectors is done by subcloning the insert cDNA into the multiple cloning sites by restriction digestion of vector and insert and maybe blunt ending with T4 DNA polymerase or Klenow polymerase, after which the fragments are ligated with T4 DNA ligase and transformed and amplified in *E. coli*. PCR is usually employed if modifications need to be made in the cDNA such as alteration of context of translation initiation or in-frame fusion to a signal sequence or another protein. Following purification of the PCR product it should be sequenced before being used for expression. The preparation of DNA is usually straightforward, but resuspension of the *E. coli* after the initial pelleting is important for efficient lysis. Centrifugation at higher speed or longer duration will make the pellet compact and difficult to resuspend. Tapping the tube on the bench aids the suspension process. One other critical step is lysis, where going beyond the recommended time is to be avoided. The quality of plasmid DNA, and especially the presence of endotoxins, has been shown to affect the transfection efficiency. If it is determined that this is an issue, a new plasmid preparation kit for endotoxin removal is available from Qiagen. The *E. coli* host chosen for plasmid preparation should be recombination deficient (*recA1* or *recA13*) and endonuclease deficient (*endA1*) to avoid rearrangement and degradation of the plasmid sequence.

Transfection

Before embarking on the laborious and lengthy process of transfection, selection, and amplification, it is advisable to do transient expression of the protein of interest. At 24 to 48 hr post transfection (which may be extended to 72 to 96 hr for secreted proteins), the cell or supernatant fraction is harvested and assayed for the protein of interest using ELISAs and immunoblots (*UNIT 10.8*). The transfection protocol may also be initially tested using reporter genes—e.g., GFP, β-galactosidase, or luciferase—in a vector construction that is identical to that of the protein of interest (Kain and Ganguly, 1996).

Lipofection (see Basic Protocol 2) and electroporation (see Alternate Protocol 1) are popular methods of gene transfer since a broad range of cell types can be transfected. The efficiency of transfection may be expressed as the total amount of gene product obtained at peak expression (24 to 48 hr), and this can be by directly assayed from cell lysates or supernatants. In cases where single-cell expression is measured using flow cytometry, as with GFP or through immunofluorescent staining, the transfection efficiency is the product of the percentage of cells positive for gene expression multiplied the average level of gene product per cell. In CHO cells, the highest transfection efficiency was obtained when 6 µL of lipid and 1 µg of DNA were used for transfecting cells in a 6-well plate (Subramanian and Srienc, 1996). The inoculation cell density and the amounts of lipid and DNA recommended in Basic Protocol 2 were obtained by scaling up in proportion to the surface area available for growth in 100-mm dishes or T-75 flasks.

The most common problems encountered in transient-expression experiments are the lack of expression and cytotoxicity. Most often, the optimal conditions are a compromise between transfection efficiency and post-pulse viability. Troubleshooting lack of expression or low-level expression can be accomplished by optimization of the conditions of transfection. In lipofection, the optimization parameters include: lipid (24 to 72 µl); DNA (4 to 16 µg); cell density; medium (αMEM, DMEM, or Opti-MEM); and incubation time with complexes (3 to 9 hr). In electroporation, the voltage is the most critical variable. In addition, the amount of DNA can be increased to 40 µg, but this increase will usually decrease the viability later. Similarly, the volume in the cuvette can be reduced by 50% or the medium could be changed to increase the conductivity. This will also increase the cytotoxicity. Cell death is more of concern in electroporation, giving viabilities after pulsing of up to 50%. This may be acceptable and may be close to optimal in

PCU IASTRI GARH UNIVERSITY LIBRARY

many cell types. In lipofection, cytotoxicity is less of a concern; here it can be reduced by decreasing the amounts of lipid and DNA or the time of incubation with complexes. The use of linearized DNA for stable transfection is recommended with the electroporation method. Circular DNA is essential for transient expression with either method. The use of necessary positive and negative controls is a must for any expression experiment. It is recommended that sensitive reporters, such as the *E. coli* β-galactosidase or the firefly luciferase, be included in transfection experiments as positive controls. Negative controls could be included through the use of no DNA or through using the vector without any protein insert. The negative control will be used to determine background (baseline) and any nonspecificity of assay.

Selection and screening

In the amplification process, after selecting the initial clones, those clones or a subset of those that represent high producers may be individually amplified. This means no further cloning is required and the danger of losing clones by outgrowth of other clones is removed. However, much more work may result, depending on the number of clones to be taken along. The amplification of individual clones is a risky enterprise since they may not amplify well. This approach may work when there are two selection markers in the same vector such as *neo* and *DHFR* (as a fusion protein or in separate transcription units) and the initial selection may be made in nucleoside-free medium containing G418. Following this initial selection of clones a few high producers may be carried on through the amplification procedure separately. The number of clones at the end of selection that need to be analyzed and the method of analysis (serum level, suspension versus adherent, production conditions, and cell growth) are very important if the requirement is to obtain a high producer. In selection and amplification, if growth is not impeded (especially in amplification) with the drug level chosen, either drug concentration should be increased or plate cell density should be decreased. In the selection procedure, if the number of cells per well exceeds what would give rise to 1 clone per well, then multiple clones will exist in the same well, which is not effective if single clones are required. The number of cells to be plated depends on the marker expression and the drug concentration. Even if the frequency of a stable transfectant is 1 in 10^5 and the number of cells plated is 100, there is always a chance that multiple clones will get into the same well, so recloning is recommended after selection/amplification procedures so as to obtain monoclonal cultures.

Limiting dilution through visual inspection is the best method, since the usual plating of 0.3 cells per well to minimize number of wells with >1 cell according to the Poisson distribution again does not assure monoclonality. Assuming Poisson distribution of viable cells, plating at a density of 0.3 cells per well would result in 74.08% wells being empty, 22.22% containing 1 cell, and ~3.69% containing >1 cell. The cloning efficiency (number of observed growth-positive wells divided by the number of expected growth-positive wells) will likely be less than expected since not all single cells form colonies as a result of stress from evaporation (higher osmolarity) or their inability to grow as single cells. Diluting the medium can be useful if evaporation and the resulting increase in osmolarity become a problem. Low serum level may reduce cloning efficiency, and it is recommended that 10% serum be included in the medium at this stage. Even at this level of stringency, it is quite possible that cultures obtained may not be monoclonal because of the growth advantage conferred upon wells with multiple cells. Visual inspection provides the only confirmation, and should be incorporated after dilution and plating to eliminate the nonsingle wells.

Almost invariably, as cells are amplified, there is no improvement in productivity after ~1 μM MTX, either because of saturation of some component in the cellular machinery or as a result of mechanisms other than amplification of DHFR for gaining resistance to methotrexate, such as a reduced influx of the drug. Amplification usually results in a high gene copy number in a nonchromosomal location, such as double minute chromosomes, which are highly unstable and which may be lost during cell division. This can be monitored as a drop in productivity following culture of cells in MTX-free medium. Low-level amplification, at or below 1 μM, is recommended to reduce instability-related issues.

Maintaining cells in culture

It is critical for any experiment that the host cells be healthy, which is made possible by maintaining them in the exponential phase of growth in medium containing 10% FBS. It is recommended that the growth of the host be checked in 10% FBS in complete medium, that

POLLACK CENTRAL LIBRARY

the cells be split at appropriate times, and that confluent cells not be exposed to prolonged nutrient limitation. A doubling time of 24 ± 8 hr is typical for most continuous cell lines. It is also advisable to start work with a frozen vial from a cell bank, which the investigator can go back to. This is important since cells that have been in culture for fewer generations (1 to 2 months after thawing) are generally easier to transfect. Some of the host cells used require special medium additives—e.g., proline for all CHO cells, as well as hypoxanthine (or adenine and deoxyadenosine), thymidine, and glycine for DHFR$^-$ CHO cells (DG44 and DXB11).

CHO cells have a tendency to clump when grown at high cell density in serum-containing medium. Clumping may be minimized by adapting cells to growth in lower serum (≤0.5% FBS). Reduced-calcium medium such as DMEM/F12 also may produce less clumping of cells. In cultures that contain clumping cells, the only secure counting method is through use of dispase and hemacytometer/trypan blue counting (*APPENDIX 3C*).

When culturing cells, it is critical to keep the temperature at 37°C ± 0.5°C. Cells will generally not tolerate rapid increases in temperature to levels above 38.5°C and will die following such shock. Exposure to lower temperature has less drastic consequences, but should be minimized so as to prevent metabolic perturbations. Optimal pH for growth is ~7.2, but cells will grow between pH 6.8 and 7.5. The optimal pH for protein expression may differ from the optimal growth temperature, and it may be worthwhile to investigate growth of cells at pH 7.0, 7.2, and 7.4 in shaker flasks to determine if there is a useful advantage to be gained in production at varied pH. All mammalian cells require L-glutamine as an energy-providing carbon source in addition to glucose. Glutaminolysis leads to the accumulation of ammonia in the medium, while glucose consumption leads to the accumulation of lactate in the medium. Both these catabolites can reach levels that may inhibit growth of cells. Typically ammonia reaches toxic levels at ~15 to 20 mM and lactate may be inhibitory at 20 mM. In cases where high cell density is desired, it is most probable that the ammonia and lactate level may limit the culture. Glutamine may also undergo spontaneous degradation to pyrrolidonecarboxylic acid and ammonia. The rate of this degradative process increases with temperature. At 37°C, ~20% of the glutamine is converted after 2 days, whereas at 4°C, <20% loss is experienced after 65 days. At small scale, medium can be purchased without glutamine and the user can add the reagent to give a concentration of 4 to 8 mM glutamine just prior to using the medium. Clearly, glutamine stock solutions (200 mM) should be stored frozen prior to use.

Ammonia and lactate can be determined rapidly using convenient metabolite analyzers such as the Kodak biolyser. In the case of large-scale operation where larger volumes of medium are needed, the user should order and use fresh complete medium with glutamine added to 4 to 8 mM. Hyperbaric levels of oxygen, which are possible where oxygen is used to supplement sparge gas, may produce peroxide and oxygen radicals which can result in poor growth and cell death. Careful control of dissolved oxygen is recommended for large-scale reactors. In the case of shaker flasks or small spinners, air is usually sufficient to support cells up to typical levels of density.

In spinners or shaker flasks, pH control is done manually, and care should be taken to add the base slowly to a swirling liquid bulk. Use of 0.5 N NaOH for pH control is recommended to avoid local high-pH spots. In reactors, use of a 6 mm internal diameter tube with a short U-bend at its end below the surface and close to the agitator is adequate. Buffering of medium using sodium bicarbonate/CO_2 is not unproblematic, and alkaline pH excursions are common when adding base or where flasks are left outside on a bench in the absence of CO_2. Such excursions may cause injure cells or affect nutrient balances. The authors recommend use of medium containing 50 to 75 mM HEPES for shaker flasks or spinner culture where manual pH adjustment is required and 25 mM in reactors where automatic pH control is available. Bicarbonate is required in the medium and is present at ~2.5 g/liter. Preincubation of spinners and flasks is important to avoid the formation of hydroxyl ions and CO_2 from the bicarbonate. Always include 5% CO_2 in the sparge and headspace gas when operating at low cell density in bioreactors. As the cell density is higher, in particular in high-density perfusion culture, the cells produce enough CO_2; therefore supplementation of sparge gas is not required. Mammalian cells are sensitive to turbulent shear, especially as may be experienced around the liquid surface where sparge gas bubbles may be bursting. Cell death can be very significant in such cases, but addition of pluronic polyol F68 (BASF) has proven to minimize bubble–cell interaction. All medium used should contain 0.05% (w/v) of Pluronic Polyol F68. Certain batches of Pluronic may be

RCU IASTECH UNIVERSITY LIBRARY

toxic (which is related to the polyoxypropylene group) and preliminary screening prior to purchase is a secure way to avoid any serious problems.

Cell growth and productivity

The specific productivity of the cell (see Support Protocols 6 and 7) is a function of the particular protein being expressed. A shaker flask growth curve (see Support Protocol 9) and titer profile will provide information on the specific productivity during the initial exponential phase and will enable calculation of the effect of lower growth rate on specific productivity as the culture progresses through to day 10. Generally the specific productivity is highest in cells growing at their maximum specific growth rate and decreases as specific growth rate is decreased. Thus, the highest specific productivity is seen in batch reactors, and high productivity can be obtained in feed-and-bleed mode. Perfusion reactors operate at lower specific growth rates with associated lower specific productivity, but provide the benefit of high cell density and higher cellular yield in the given medium. In cases where there is evidence of proteolysis or glycosidase degradation of the product, it is advisable to operate a perfusion system with high viability. In this type of system the product residence time at 37°C and cell lysis are minimized. Further minimization of proteolysis may be obtained by addition of specific protease inhibitors such as aprotinin (for inhibition of serine proteases). Addition of sialidase inhibitors such as 2,3-dehydro-2-deoxy-*N*-acetylneuramic acid (Gramer et al., 1995) has been described. Many of the low-molecular-weight inhibitors are toxic and undesirable in production of material destined for clinical applications. In such cases, rapid processing at low temperature (<10°C) is the best option. Where proteins are exquisitely sensitive to proteolysis, it is advisable to operate with 0.5% to 1.0% FBS in the cell culture medium as a reagent to prevent proteolysis (Tiege et al., 1994), as serum contains the potent protease inhibitor α2-macroglobulin.

Many cell culture–dictated conditions may influence the microheterogeneity of the glycoprotein produced. In order to reproduce equivalent product from cell culture batches, it is important to standardize the cell culture protocol.

Spin filters are a commonplace option for cell retention in bioreactors; however, fouling of the screen with cells is usual after 2 to 3 weeks of operation. The passive settler devices are low-maintenance and predictable for cells of a given size and clumpiness. At present, settlers are custom-designed, although several approaches have been described in the literature (Batt et al., 1990; Brown et al., 1993).

Harvesting cultures

When harvesting cultures, it is important to maintain the cells intact to prevent contamination of the product in the extracellular compartment with intracellular constituents or to prevent loss of product if it is intracellular. Use of peristaltic pumps with size 16 or 24 silicone tubing or use of gravity transfer is recommended. Static laboratory centrifugation is a gentle approach and useful for clarifying relatively small volumes of culture liquid. The largest rotor capacities are ~5 liters, and the researcher can gauge the time investment needed to process the material with such an approach. It would be feasible to process up to 30 liters of culture broth in ~2.5 hr using such an approach.

Where cross-flow filtration is employed to separate cells and medium, it is important to minimize cell lysis, which could result in a decrease in yield, purity, or performance of the membrane system. Operation at cold-room temperature (4° to 12°C) is advisable as in all of the downstream recovery operations. Avoidance of average wall shear rates greater than 3000 sec^{-1}, using membrane pore sizes of 0.2 μm and minimization of pump shear rates will minimize the potential for cell damage (Maiorella et al., 1991). If small membrane systems of 1 ft^2 are used, then peristaltic pumps (Watson Marlow or Cole-Parmer) are adequate, while for larger systems the larger-gap rotary positive- displacement pumps are employed (Flowtech, Waukesha). Cross-linked cellulosic membranes are available (Sartorius Hydrosart) which are useful in minimizing surface fouling and for facilitating high protein flux.

Anticipated Results

For plasmid preparation, 1 liter of *E. coli* culture grown to an OD_{650} of 4.0 can yield up to 5 mg of high-copy-number plasmids. The DNA may be stored frozen without loss of quality.

Confluent growth on a surface corresponds to a cell-surface density of 1–1.5 × 10^5 cells/cm^2, so one T-75 flask can provide ~1 × 10^7 total viable cells and one T-175 flask can provide ~2–3 × 10^7 total viable cells, enough for the transfection process. Approximately 50% of the cells will have taken up the plasmid

BOURKE CENTRAL LIBRARY

Table 5.10.2 Specific Productivities for a Selection of Proteins Expressed in Chinese Hamster Ovary Cells

Protein	Cell Line	Promoters		Amplification	Productivity ($\mu g/10^{-6}$ cells/day)	Notes	Reference
		GOI	SM				
Heterologous proteins							
TIMP	CHO-K1	hCMV-TIMP	SV40 late-GS	20 μM MSX	110	copy number ~30	Cockett (1990)
TIMP	CHO DXB11	hCMV-TIMP	SV40 early-dhfr	1 μM MTX	17.5	copy number~100,	
htPA	CHO DXB11	Ad2 MLP-tPA	Ad2 MLP-dhfr	0.5 μM MTX	66	dual plasmid cotransfection	Kaufman (1985)
IFN-β	CHO DXB11	SV40 early-IFN	SV40 early-dhfr	1 μM MTX	0.5	dual plasmid cotransfection	Chernajovsky (1984)
IFN-γ	CHO DXB11	SV40 early-IFN	SV40 early-dhfr	0.3 μM MTX	10	dicistronic vector	Mory (1966)
h-AT III	CHO DXB11	SV40 early-ATIII	MMTV-LTR-dhfr	10 μM MTX	22		Zettlmeissl (1987)
Monoclonal antibodies							
Campath-1H	CHO DXB11	β-actin–light chain β-actin–heavy chain	weak SV40-dhfr MMT-neo	1 μM MTX	100	Dual plasmid transfection	Page (1991)
Subunit vaccines							
HBsAg	CHO DHFR$^-$	SV40 early-HBV	MMTV-LR dhfr	10 μM MTX	1-2 (stable)	copy number ~400 (T-flask)	Michel (1985)

[a]Abbreviations: GOI, gene of interest; MTX, methotrexate; SM, selection marker.

vector following lipofection versus approximately 10% in electroporation. The proportion of cells that chromosomally integrate the foreign DNA is low, in the range of 0.001% to 0.01%. Of these cells, the frequency of stable high producers ~0.1% in general, but may reach 1% in cases where targeted integration methods are used. From these numbers, it is apparent that an initial total of 1×10^7 cells will give ~10 to 100 productive stable clones with perhaps 1 to 10 clones being high producers.

The specific productivity of the transfected CHO cells will be heterogeneous within the cell population. The level and range of productivity is largely a function of the protein being expressed as well as the particular culture conditions. Typical specific productivity levels are ~5 to 20 $\mu g/10^6$ cells/day, although it may be higher, with examples indicating recombinant antibody production at >100 $\mu g/10^6$ cells/day. Lower values, ~1 to 5 $\mu g/10^6$ cells/day, are typically found for recombinant viral-antigen proteins.

Amplification may increase the specific productivity 10- to 100-fold, depending on the protein and the clone selected. Examples of productivities following amplification that have been reported in the literature are provided in Table 5.10.2. Where a high producer is initially selected, the amplification process may provide little gain in productivity. Productivity may be limited by secretion rates in the case of large glycoproteins with numerous disulfides.

CHO cells adapt quickly to growth in suspension and can be adapted to growth in low (0.5%) serum in ~15 to 20 passages. High specific growth rate is usually reestablished in 3 to 5 passages following a step change in a medium concentration. Complete elimination of serum may take a further 5 to 10 passages before cells become truly serum-independent. Once cells are adapted to growth in medium such as DMEM/F12, it is possible to obtain a cell density of $2–4 \times 10^6$ per ml in batch shaker flasks or spinners. In the case of bioreactors, with precise pH and dissolved oxygen control, cell densities may reach 4×10^6 cells/ml for batch operation and as high as 10×10^6 cells/ml for bioreactors operating in perfusion mode. Product titers of 5 to 20 mg/liter are possible in batch-mode flasks, spinners, and bioreactors. Fed batch reactors may provide higher titers,

ECU IASTATE UNIVERSITY LIBRARY

since the promoters that have been recommended are generally dominant in the S phase of the CHO cell cycle and faster specific growth will result in a higher specific productivity than slower growing cells. A cell that has a specific productivity of 5 pg/cell/day at a specific growth rate of 0.3 per days is capable of providing 75 to 100 mg/day/liter culture in a perfusion reactor operating with a cell density of 15×10^6 cells/ml, a dilution rate of 1.0, and 70% overall cell retention.

The purity of the product, with a titer of 30 mg/liter, is estimated at ~5% for cells grown in bioreactors that achieve 10×10^6 cells/ml with a viability of 90%, using 1% FBS. In serum-free medium under the same conditions, the purity may be 10%. Intracellular-produced proteins rarely exceed 1% of total intracellular protein.

Recovery yield of the product from the reactor should be in the region of 75% to 90%. At large scale, the recovery yield for intracellular protein may be lower (50% to 60%) if large volumes of high-debris-containing material result in membrane fouling. Use of cellulose-based membranes and/or depth filtration may improve such low yields.

Time Considerations

Preparation of DNA and the process of transfection can be accomplished in <1 week. Following transfection, it takes 3 weeks to obtain colonies on plates. Expansion of the colonies up to the tissue culture flask stage requires 2 further weeks. Thus it will take a total of ~6 weeks to obtain stable transfected cells that grow in T-75 flasks in the presence of the selection drug. Subjecting the cells to amplification, such as an increased level of methotrexate, will take up 3 weeks to produce a pool of cells. Each further round of amplification will require 3 weeks. Based on these time lines, a cell that has been subjected to two rounds of amplification would be available 12 weeks after initiating the full process.

Following amplification, the process of single-cell cloning requires 3 weeks to obtain colonies and a further 2 weeks to scale up to the T-75 flask stage. At this point the higher-producing clones can be adapted from tissue culture flask growth to shaker flask growth in 1 week, and can be further adapted to 1% serum in a 2 weeks, 0.5% serum in 5 weeks, and 0% serum in 8 weeks. Laying down a cell bank of the adapted cells takes ~1.5 weeks.

Based on a cell-growth doubling rate of 1 to 2 days, a batch culture requires 5 to 7 days to complete, and depending on the scale would require an inoculum buildup process. For a 100-liter-scale process it may take 10 to 15 days to scale up to a suitable inoculum volume.

Recovery of product is usually performed as a batch process and should require <8 to 10 hr.

Literature Cited

Aitken, A. 1997. A library of concensus sequences. *Methods Mol. Biol.* 64:357-367.

Aruffo, A. 1997. Transient expression of proteins in COS cells. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.13.1-16.13.7. John Wiley & Sons, New York.

Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., (eds.) 1998. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Batt, B.C., Davis, R.H., and Kompala, D.S. 1990. Inclined sedimentation for selective retention of viable hybridomas in a continuous suspension bioreactor. *Biotechnol. Prog.* 6:458-464.

Brady, R.O., Murray, G.J., and Barton, N.W. 1994. Modifying exogenous glucocerebrosidase for effective replacement therapy in Gaucher Disease. *J. Inherited Metab. Dis.* 17:510-519.

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E., and Palmiter, R.D. 1988. Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 85:836-840.

Brown, P.C., Wininger, M., Chow, R., and Cox, J. 1993. Perfusion culture of CHO cells in suspension facilitated by inclined sedimentation chambers for cell recycle. Presented at the 205th National Meeting of the ACS, Division of Biotechnology, March 31, 1993.

Budelier, K. and Schorr, J. 1998. Purification of DNA by anion-exchange chromatography. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 2.1.11-2.1.18. John Wiley & Sons, New York.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. 1994. Green fluorescent protein as a marker of gene expression. *Science* 263:802-805.

Chalmers, J.J., Zborowski, M., Sun, L., and Moore, L. 1998. Flow- through immunomagnetic cell separation. *Biotechnol. Prog.* 14:141-148.

Chernajovsky, Y., Mory, Y., Chen, L., Marks, Z., Novick, D., Rubinstein, M., and Revel, M. 1984. Efficient constitutive production of himan fibroblast interferon by hamster cells transformed with the IFN-β1 gene fused to an SV40 early promoter. *DNA* 3:297-308.

Chesnut, J.D., Baytan, A.R., Russel, M., Chang, M-P., Bernard, A., Maxwell, I.H., and Hoeffler, J.P. 1996. Selective isolation of transiently transfected cells from a mammalian cell population with vectors expressing a membrane anchored

RESEARCH CENTRAL UNIVERSITY LIBRARY

single-chain antibody. *J. Immunol. Methods* 193:17-27.

Cockett, M.I., Bebbington, C.R., and Yarranton, G.T. 1990. High level expression of tissue inhibitor of metaloproteinase in Chinese hamster ovary cells using glutamine synthetase gene amplification. *Biotechnol.* 8:662-667.

Earl, P. and Moss, B. 1998. Characterization of recombinant vaccinia viruses and their products. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.18.1-16.18.10. John Wiley & Sons, New York.

Earl, P., Wyatt, L., Carroll, M., and Moss, B. 1998a. Generation of recombinant vaccinia viruses. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.17.1-16.17.16. John Wiley & Sons, New York.

Earl, P.L., Cooper, N., Wyatt, L., Carroll, M., and Moss, B. 1998b. Preparation of cell cultures and vaccinia virus stocks. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.16.1-16.16.7. John Wiley & Sons, New York.

Elroy-Stein, O. and Moss, B. 1998. Gene expression using the vaccinia virus/T7 RNA polymerase hybrid system. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.19.1-16.19.9. John Wiley & Sons, New York.

Fukushige, S. and Sauer, B. 1992. Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 89:7905-7909.

Gramer, M.J., Goochee, C.F., Chock, V.Y., Brousseau, D.T., and Sliwkowski, M.B. 1995. Removal of sialic acid from a glycoprotein in CHO cell culture supernatant by action of an extracellular CHO cell sialidase. *Bio/Technology* 13:692-298.

Gray, F., Kenney, J.S., and Dunne, J.F. 1995. Secretion capture and report web: Use of affinity derivatized agarose microdroplets for the selection of hybridoma cells. *J. Immunol. Methods* 182:155-163.

Griffiths, B. 1992. Scaling-up of animal cell cultures. *In* Animal Cell Culture, 2nd ed. (R.I. Freshney, ed.), pp. 47-93. IRL Press, Oxford.

Gu, M.B., Todd, P., and Kompala, D.S. 1993. Foreign gene expression (β-galactosidase) during the cell cycle phases in recombinant CHO cells. *Biotechnol. Bioeng.* 42:1113-1123.

Kahana, J. and Silver, P. 1996. Use of the *A. victoria* green fluorescent protein to study protein dynamics in vivo. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 9.7.22-9.7.28. John Wiley & Sons, New York.

Kain, S. and Ganguly, S. 1996. Overview of genetic reporter systems. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 9.6.3-9.6.12. John Wiley & Sons, New York.

Kaufman, R.J. 1978. Quantitation of dihydrofolate reductase in individual parental and methotrexate-resistant murine cells. *J. Biol. Chem.* 16:5852-5860.

Kaufman, R.J. 1990. Selection and coamplification of heterologous genes in mammalian cells. *Methods Enymol.* 185:537-566.

Kaufman, R.J., Wasley, L.C., Spiliotes, A.J., Gossels, S.D., Latt, S.A., Larsen, G.R., and Kay, R.M. 1985. Coamplificaton and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in Chinese hamster ovary cells. *Mol. Cell. Biol.* 5(7):1750-1759.

Kaufman, R.J., Davies, M.V., Wasley, L.C., and Michnick, D. 1991. Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus. *Nucl. Acids Res.* 19(16):4485-4490.

Kingston, R.E., Kaufman, R.J., Bebbington, C.R., and Rolfe, M.R. 1993. Amplification using CHO cell expression vectors. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.), pp. 16.14.1-16.14.13. John Wiley & Sons, New York.

Klickstein, L.B. and Neve, R.L. 1991. Ligation of linkers or adapters to double-stranded cDNA. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.), pp. 5.6.1-5.5.10. John Wiley & Sons, New York.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44:283-292.

Kriegler M. 1990. Gene Transfer and Expression, A Laboratory Manual. Stockton Press, New York.

Liljestrom, P. and Garoff, H. 1995. Expression of proteins using Semliki Forest virus vectors. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.20.1-16.20.16. John Wiley & Sons, New York.

Lucas, B.K., Giere, L.M., DeMarco, R.A., Shen, A., Chisholm, V., and Crowley, C.W. 1996. High level expression of recombinant proteins in CHO cells using adicistronic *DHFR* intron expression vector. *Nucl. Acids Res.* 24(9):1774-1779.

Maiorella, B.M., Dorin, G., Carion, A., and Harano, D. 1991. Crossflow microfiltration of animal cells. *Biotechnol. Bioeng.* 37:121-126.

FCU IASI LIBRARY

Michel, M-L., Sobczak, E., Malpiece, Y., Tiollais, P., and Streeck, R.E. 1985. Expression of amplified hepatitis B surface antigen in Chinese hamster ovary cells. *Bio/Technology* 3:561-566.

Miloux, B. and Lupker, J.H. 1994. Rapid isolation of highly productive recombinant Chinese hamster ovary cell lines. *Gene* 149:341-344.

Morris, A.E., Lee, C-C., Hodges, K., Aldrich, T.L., Krantz, C., Smidt, P.S., and Thomas, J.N. 1997. Expression augmenting sequence element (EASE) isolated from Chinese hamster ovary cells. *In* Animal Cell Technology (M.J.T. Carrondo, ed.) pp. 529-534. Kluwer Academic Publishers, Boston.

Mortensen, R., Chesnut, J.D., Hoeffler, J.P., and Kingston, R.E. 1997. Selection of transfected mammalian cells. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D.Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.5.1-16.5.19. John Wiley & Sons, New York.

Mory, Y., Ben-Barak, J., Segev, D., Cohen, D., Novick, D., Fischer, D.G., Rubinstein, M., Kargman, S., Zilberstein, M., and Revel, M. 1986. Efficient constitutive production of human IFN-γ in Chinese hamster ovary cells. *DNA* 5:181-193.

Moss, B. and Earl, P.L. 1998. Overview of the vaccinia virus expression system. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.15.1-1615.5. John Wiley & Sons, New York.

Page, M.J. and Sydenham, M.A. 1991. High level expression of the humanized monoclonal antibody Campath-1H in Chinese hamster ovary cells. *Bio/Technology* 9:64-68.

Parekh, R.B. and Patel, T.P. 1992. Comparing the glycosylation patterns of recombinant glycoproteins. *Trends Biotechnol.* 10(8):276-280.

Pendse, G.J., Karkare, S., and Bailey, J.E. 1992. Effect of cloned gene dosage on cell growth and Hepatitis B surface antigen synthesis and secretion in recombinant CHO cells. *Biotechnol. Bioeng.* 40:119-129.

Robinson, J.P., Darzynkiewicz, Z., Dean, P.N., Orfao, A., Rabinovitch, P.S., Stewart, C.C., Tanke, H.J., and Wheeless, L.L. (eds.) 1998. Current Protocols in Cytometry. John Wiley & Sons, New York.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Simonsen, C. and Levinson, A. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 80:2495-2499.

Stanley, P. 1989. Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol. Cell Biol.* 9:377-383.

Stanley, P. and Ioffe, E. 1995. Glycosyltransferase mutants: Key to new insights in glycobiology. *FASEB J.* 9:1436-1444.

Subramanian, S. and Srienc, F. 1996. Quantitative analysis of transient gene expression in mammalian cells using green fluorescent protein. *J. Biotechnol.* 49:137-151.

Tiege, M., Weidemann, and R., Kretzmer, G. 1994. Problems with serum-free production of antithrombin III regarding proteolytic activity and product quality. *J. Biotechnol.* 34:101-105.

Wurm, F.M., Pallavicini, M.G., and Arathoon, R. 1992. Integration and stability of CHO amplicons containing plasmid sequences. *Dev. Biol. Stand.* 76:69-82.

Wurm, F.M., Johnson, A., Ryll, T., Kohne, C., Scherthan, H., Glaab, F., Lie, Y.S., Petropoulos, C.J., and Arathoon, W.R. 1996. Gene transfer and amplification in CHO cells: Efficient methods for maximizing specific productivity and assessment of genetic consequences. *Ann. N.Y. Acad. Sci.* 782:70-78.

Yenofsky, R.L., Fine, M., and Pellow, J.W. 1990. A mutant phosphotransferase II gene reduces the resistance of transformants to antibiotic selection pressure. *Proc. Natl. Acad. Sci. U.S.A.* 87:3435-3439.

Yim, K.W. 1991. Fractionation of human recombinant tissue plasminogen activator (rtPA) glycoforms by high-performance capillary zone electrophoresis and capillary isoelectric focusing. *J. Chromatogr.* 559:401-410.

Zettlmeissel, G., Ragg, H., and Karges, H.E. 1987. Expression of biologically active Antithrombin III in Chinese hamster ovary cells. *Bio/Technology* 5:720-725.

Key References

Butler, M. (ed.). 1991. Mammalian Cell Biotechnology, IRL Press, Oxford.

Excellent introductory background with protocol and tips.

Freshney R.I. (ed.). 1992. Animal Cell Culture. (2nd ed.), IRL Press, Oxford.

Practical guide to cell culture with useful protocols and tips.

Kriegler, 1990. See above

Comprehensive manual for molecular biology.

Contributed by Su Chen, David Gray, Johnny Ma and Shyamsundar Subramanian
Chiron Corporation
Emeryville, California

BCU/CENTRAL UNIVERSITY LIBRARY

PURIFICATION OF RECOMBINANT PROTEINS

6

INTRODUCTION 6.0.1

6.1 Overview of the Purification of Recombinant Proteins Produced in *Escherichia coli* **6.1.1**
- Protein Sequence and Composition Analysis 6.1.1
- Characteristics of the Host-Vector System 6.1.2
- Solubility and Location of the Protein 6.1.4
- Strategies for Isolation of Soluble Proteins 6.1.7
- Strategies for Isolation of Insoluble Proteins 6.1.11
- Bacterial Expression of Proteins Normally Glycosylated 6.1.15
- Protein Handling 6.1.16
- Scale of Operations and Aims of Purification 6.1.17
- Specialized Equipment 6.1.18
- Literature Cited 6.1.19

6.2 Preparation of Soluble Proteins from *Escherichia coli* **6.2.1**
- Basic Protocol: Purification of a Protein Expressed in *Escherichia coli* in a Soluble State: Interleukin 1β 6.2.1
- Reagents and Solutions 6.2.9
- Commentary 6.2.10

6.3 Preparation and Extraction of Insoluble (Inclusion-Body) Proteins from *Escherichia coli* **6.3.1**
- Basic Protocol 1: Preparation and Extraction of Insoluble (Inclusion-Body) Proteins from *Escherichia coli* 6.3.1
- Basic Protocol 2: Medium-Pressure Gel-Filtration Chromatography in the Presence of Guanidine Hydrochloride 6.3.3
- Support Protocol: Preparation of Samples Containing Guanidine Hydrochloride for SDS-PAGE 6.3.6
- Reagents and Solutions 6.3.7
- Commentary 6.3.8

6.4 Overview of Protein Folding **6.4.1**
- How Proteins Fold 6.4.1
- How to Fold Proteins 6.4.2
- Literature Cited 6.4.7

6.5 Folding and Purification of Insoluble (Inclusion-Body) Proteins from *Escherichia coli* **6.5.1**
- Basic Protocol 1: Folding and Purification of Bovine Growth Hormone 6.5.2
- Basic Protocol 2: Folding and Purification of Human Interleukin 2 6.5.6
- Support Protocol: Resolution of Native and Misfolded Forms of hIL-2 by RP-HPLC 6.5.8
- Basic Protocol 3: Folding and Purification of a Histidine-Tagged Protein: HIV-1 Integrase 6.5.10
- Reagents and Solutions 6.5.14
- Commentary 6.5.18

continued

BCU IAS / CENTRAL UNIVERSITY LIBRARY

Copyright © 1997 by John Wiley & Sons, Inc.

6.6 Expression and Purification of GST Fusion Proteins **6.6.1**
Strategic Planning 6.6.1
Basic Protocol 1: Expression of Glutathione-*S*-Transferase Fusion Protein 6.6.5
Basic Protocol 2: Affinity Chromatography Purification of a Soluble GST Fusion Protein 6.6.8
Alternate Protocol 1: Affinity Chromatography Purification of GST Fusion Protein from Inclusion Bodies 6.6.10
Alternate Protocol 2: Batch Purification of GST Fusion Protein 6.6.11
Basic Protocol 3: Protease Cleavage of Fusion Protein in Solution to Remove GST Affinity Tag 6.6.12
Alternate Protocol 3: Protease Cleavage of Fusion Protein Bound to a Glutathione-Sepharose Column 6.6.14
Alternate Protocol 4: Batch Protease Cleavage and Separation of Fusio Protein on Glutathione Resin 6.6.15
Support Protocol 1: Affinity Chromatography Purification of Polypeptides After Enzymatic Cleavage 6.6.15
Support Protocol 2: Further Purification of Cleaved Recombinant Protein Using HPLC Gel Filtration 6.6.16
Reagents and Solutions 6.6.17
Commentary 6.6.19

6.7 Expression and Purification of Thioredoxin Fusion Proteins **6.7.1**
Strategic Planning 6.7.1
Basic Protocol: Construction and Expression of a Thioredoxin Fusion Protein 6.7.4
Support Protocol 1: *E. coli* Lysis Using a French Pressure Cell 6.7.6
Support Protocol 2: Osmotic Release of Thioredoxin Fusion Proteins 6.7.7
Support Protocol 3: Purification of Thioredoxin Fusion Proteins by Heat Treatment 6.7.8
Reagents and Solutions 6.7.9
Commentary 6.7.10

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 6
Purification of Recombinant Proteins

INTRODUCTION

The principle aim of protein overexpression using the various systems described in Chapter 5 and elsewhere (e.g., Chapter 16 of Ausubel et al., 1997) is to obtain purified protein, and that is the subject matter of this chapter. The purification of any protein, recombinant or nonrecombinant, is critically dependent on the starting material (or source). Some natural proteins—for example, cellular receptors and growth factors—may be present in biological source material only in trace amounts. In these cases, complex purification schemes are required to obtain micro-amounts of partially or completely pure protein (Simpson and Nice, 1989, and references therein). With recombinant sources, the overproduced protein is usually (though not always) abundant, and hence the amount available is not generally a critical factor. The gene expressing the recombinant protein must of course have been previously cloned, and in fact may have been done so using oligonucleotide probes derived from the amino acid sequence data, in turn derived from the authentic protein isolated by (to use Scope's expression) "heroic efforts" (*UNIT 1.1*). What is at issue with overproduced proteins made by the various host/vector systems is their authenticity, in terms of both noncovalent (i.e., conformational—is the protein correctly folded or aggregated?) and covalent structure (i.e., is the correct disulfide bond pattern present and are post-translational modifications such as glycosylation required?).

In *UNIT 6.1*, an overview of protein purification is presented; although purification from *E. coli* is emphasized, other host systems are also referred to where appropriate. Purification of recombinant protein can be aided by the fact that the investigator knows the primary sequence of the recombinant protein and can thus estimate its isoelectric point. Furthermore, the purified protein can thus be checked against this theoretical sequence for artificial modifications arising from the expression system and purification method using some of the methods summarized in Chapter 7.

If the recombinant source contains the protein to be purified in a soluble state, then the protein can be isolated using standard methods, often involving column chromatography. In *UNIT 6.2*, the purification of interleukin 1β (IL-1β) is used to illustrate the basic approach for purifying an abundant and soluble protein expressed in *E. coli*. Using the basic fractionation and chromatographic approach described, it is often possible to purify other soluble recombinant proteins. For situations where the starting material is less abundant, albeit soluble (e.g., proteins secreted into the medium, or, in the case of *E. coli*, the periplasmic space), higher-resolution techniques such as biospecific affinity chromatography may be required. These are discussed in *UNIT 6.1* and Chapter 9. The expression of proteins genetically engineered to contain various affinity tags or carrier proteins that aid purification are reviewed in *UNIT 5.1* as well as in *UNITS 6.1, 6.5, 6.6 & 6.7* of this chapter.

IL-1β is somewhat of an ideal protein from the viewpoint of both expression and purification. The authentic protein is not post-translationally modified and although it contains two sulfhydryl groups, it contains no disulfide bonds. For this protein, the authentic and recombinant forms have the same covalent structure, as is discussed in detail in *UNIT 6.2* (Background Information).

Contributed by Paul T. Wingfield
Current Protocols in Protein Science (1997) 6.0.1-6.0.4
Copyright © 1997 by John Wiley & Sons, Inc.

BOND UNIVERSITY LIBRARY

In many cases high-level expression of protein in the *E. coli* host leads to the formation of highly aggregated proteins, called inclusion bodies. The determination of whether a protein is soluble or not is made using cell lysates and simple centrifugation schemes as outlined in *UNIT 6.1* (see also *UNIT 5.3*). Proteins in inclusion bodies usually have the correct primary sequences, identical to those of their native and authentic counterparts, but are aggregated and inactive due to noncovalent or conformational differences. In *UNIT 6.3*, the preparation of insoluble inclusion body protein from *E. coli* lysates by a simple washing procedure is described.

Before insoluble proteins in inclusion bodies can be purified, they must be extracted and solubilized; in this respect their purification is analogous to the extraction of intrinsic membrane proteins from membranes. The solubilization of inclusion body protein, however, requires denaturation with reagents such as concentrated solutions of urea or guanidine·HCl (see *APPENDIX 3A*), whereas the extraction of membrane proteins is normally carried out under nondenaturing conditions using nonionic or zwitterionic detergents (von Jagow et al., 1994). As the solubilized recombinant protein is denatured or unfolded, it must be folded into the correct tertiary and quaternary structures in order for the protein to acquire its functional and biological properties.

Many nonrecombinant proteins that have been isolated under native conditions can be reversibly denatured and renatured using well-established methods. A concise account of protein folding is given in *UNIT 6.4* (see also Pain, 1994) to provide some of the theoretical background required to interpret protein folding experiments.

The preparation of folded and active recombinant proteins from inclusion bodies involves the process known as "preparative protein folding." In traditional protein folding experiments, pure native proteins such as ribonuclease were used to study the reversible denaturation/renaturation process to gain information about folding pathways (Pain, 1994). These experiments are usually carried out using low protein concentrations (in the micrograms per milliliter range) to minimize nonspecific inter- and intramolecular interactions that lead to aggregation. In preparative protein folding, the aim is to obtain folded protein in as high a yield as possible using relatively high starting concentrations (in the milligrams per milliliter range) and without using very large reaction volumes. The various strategies usually used to achieve this are illustrated in *UNIT 6.5*.

In preparative protein folding, apart from the need to avoid or minimize aggregation, formation of the correct disulfide bonds is required. After extraction of the recombinant protein from inclusion bodies using a protein denaturant and reducing agent (*UNIT 6.3*), the solubilized protein is unfolded and sulfhydryl groups, if present, are reduced. To form the correctly folded protein, disulfide bond formation and protein folding must both occur, and usually in a concomitant manner.

Folding pure or partially purified protein may be advantageous, leading to higher yields of folded protein. In *UNIT 6.3*, gel filtration in the presence of guanidine·HCl is described as a simple and rapid method for partially purifying extracted and denatured proteins prior to folding. A recent example of this approach is the production of the ectodomain of HIV-1 and SIV gp41 (Wingfield et al., 1997). Other methods that can be used to purify proteins in high concentrations of urea or guanidine·HCl are discussed in *UNIT 6.1*.

Three examples of the folding and purification of insoluble proteins expressed in *E. coli* are described in *UNIT 6.5*. Bovine growth hormone (somatotropin) is extracted from inclusion bodies using guanidine·HCl; the protein is then folded and then purified by conventional chromatographic methods. Aggregation during the folding process is minimized by using a solvent additive cosolvent, in this case urea in a limiting (i.e., nonde-

BCU IAS / CENTRAL UNIVERSITY LIBRARY

naturing) concentration. Oxidation of the two disulfide bonds is enhanced by inclusion of an oxido-shuffling system based on oxidized and reduced glutathione (*UNITS 6.1 & 6.4*).

In the second example, purification of human interleukin-2 (T cell growth factor), a different approach is used. The protein is extracted from inclusion bodies using acetic acid, then fractionated by gel filtration in the same acid. The partially purified protein is folded and air-oxidized by the simple approach of dialysis against water. Because interleukin-2 contains three sulfhydryl groups, only two of which are involved in an intramolecular disulfide bond, it is possible for incorrect disulfide bonds to form during folding. An HPLC-based support protocol is included that allows rapid determination of the disulfide bonding pattern. Finally, the folded protein is purified to homogeneity using preparative reversed-phase HPLC. The acid stability of the protein is clearly an advantage and necessary for the success of the overall process.

The third example describes the purification of the catalytic domain of the immunodeficiency virus (HIV-1) integrase. This protein is essential for the HIV-1 life cycle and is therefore a potential drug target. The integrase catalytic domain is expressed as a fusion protein containing an additional 20 amino acids at the N-terminus. This "fusion domain" includes a tract of six histidine residues and is commonly called a His tag. The expression of proteins with His tags is fairly common practice, as it allows proteins to be purified by metal-chelate chromatography (MCAC) under both native and denaturing conditions (see *UNIT 5.1* for additional discussion). The HIV-1 integrase, extracted with denaturant and partially purified by gel filtration and MCAC, is folded into native protein using a slow and continuous addition of denatured protein into a renaturation solution. This approach was originally described by Ruldolph and his colleagues (see Ruldolph, 1989, and references therein) and is generally applicable for avoiding and/or minimizing aggregation during folding. The His tag is removed from the folded fusion protein using specific protease digestion (in this example, the protease used is thrombin, and methods for its handling and specific removal are also described).

A widely used method for producing fusion proteins in *E. coli* is the glutathione *S*-transferase (GST) system (*UNIT 5.1*). Analogous to His tag fusions, there is a simple affinity chromatographic method for purification of GST fusions using immobilized glutathione as an affinity matrix. The affinity purification of GST fusions is probably more specific than the metal chelate chromatography used for His tags, but it has the limitation that it cannot be performed under denaturing conditions. Fusion proteins usually contain a linker region with a specific proteolytic cleavage site, and often this site is specific for thrombin. Thus, once the GST fusion protein has been purified, the protein moiety can be readily released by proteolysis. Further purification may include one more round of affinity chromatography (to remove GST) and a gel filtration step to remove excess thrombin and other minor contaminates. Purification may be possible by performing only the gel filtration step, depending on the size of various protein products (GST is a dimer of 51 kDa and thrombin is ~36 kDa). In *UNIT 6.6* the expression and purification of GST fusion proteins is described. Protocols for protein expression using shaker cultures and purification methods of soluble GST fusion proteins are included. If the GST fusion protein is insoluble and forms inclusion bodies, the protein must be solubilized and folded into a native-like conformation prior to affinity chromatography. An extraction and folding protocol using the protein denaturant urea is described. (It should be noted that the GST moiety can be readily folded from either urea or guanidine·HCl denatured samples.) Protocols for proteolytic cleavage of the fusion protein and further purification are also described. Structural biologists may note with interest the recent paper by Kuge et al. (1997) in which the crystallization of a GST fusion protein was achieved.

BOWLING GREEN STATE UNIVERSITY LIBRARY

Another popular gene fusion system is based on the *E. coli* intracellular enzyme thioredoxin. This 12-kDa protein is very soluble, is monomeric, and has been thoroughly characterized (*UNIT 5.1*). The target protein can be positioned either N-terminal (protein-thioredoxin) or C-terminal (thioredoxin-protein) to the fusion partner. This dual orientation, also possible with Hist-tag fusions, represents a difference from GST fusions, as the GST moiety is normally only N-terminal. It is worth noting that as thioredoxin is a monomeric protein, it may be a more suitable fusion partner for oligomeric proteins than, for example, dimeric GST (Hurd and Hornby, 1996). In *UNIT 6.7* the construction and expression of thioredoxin fusion proteins are discussed. Support protocols deal with the release of the fusion proteins by mechanical lysis and osmotic stress. *E. coli*–derived thioredoxin exhibits thermal stability, and a heat treatment treatment of cell lysates at 80°C may be a useful initial purification step. This approach is described in the third support protocol. The commentary section includes a brief discussion of more specific purification methods. However, if ease of purification is of critical importance, the GST fusion system should be considered as the glutathione-based affinity method is probably more efficient than the corresponding affinity-based method used for thioredoxin.

It should be emphasized that although most of the protocols described in this chapter are applied to specific examples, they are fairly representative of the methods and approaches used in general for laboratory-scale preparation of recombinant proteins from *E. coli*. The individual characteristics of a given protein (e.g., solubility, isoelectric point, pH stability, and number of sulfhydryl residues) will dictate the conditions required for any particular purification step or folding process. Method optimization is usually empirical unless characterization of the purified protein provides useful guidelines.

LITERATURE CITED

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K.S. (eds.). 1997. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Hurd, P.J. and Hornby, D.P. 1996. Expression systems and fusion proteins. *In* Proteins LabFax (N.C. Price, ed.) pp. 109-117. BIOS Scientific Publishers, Oxford, UK.

Kuge, M., Fujii, Y., Shimizu, T., Hirose, F., Matsukage, A., and Hakoshima, T. 1997. Use of a fusion protein to obtain crystals suitable for X-ray analysis: Crystallization of a GST-fusion protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF. *Protein Sci.* 6:1783-1786.

Pain, R.H. 1994. Mechanisms of Protein Folding. IRL Press, Oxford, UK.

Ruldolph, R. 1989. Renaturation of recombinant, disulfide-bonded proteins from "inclusion bodies." *In* Modern Methods in Protein and Nucleic Acid Research: Review Articles (H. Tschesche, ed.) pp. 149-171. Walter de Gruyter, Berlin.

Simpson, R.J. and Nice, E.C. 1989. Strategies for the purification of subnanomole amounts of proteins and polypeptides for microsequence analysis. *In* The Use of HPLC in Receptor Biochemistry (A.R. Kerlavage, ed.) pp. 210-244. Alan R. Liss, New York.

von Jagow, G., Link, T.A., and Schägger, H. 1994. Purification strategies for membrane proteins. *In* A Practical Guide to Membrane Protein Purification (G. von Jagow and H. Schägger, eds.) pp. 3-21. Academic Press, San Diego.

Wingfield, P.T., Stahl, S.J., Kaufman, J., Zlotnick, A., Hyde, C.C., Gronenborn, A.M., and Clore, G.M. 1997. The extracellular domain of immunodeficiency virus gp41 protein: Expression in *Escherichia coli*, purification, and crystallization. *Protein Sci.* 6:1653-1660.

Paul T. Wingfield

BCU IAS CENTRAL UNIVERSITY LIBRARY

Overview of the Purification of Recombinant Proteins Produced in *Escherichia coli*

The initial choice of host system for the production of recombinant proteins for many investigators is *Escherichia coli*. This is due to such factors as ease of genetic manipulation, availability of optimized expression plasmids, and ease of growth. This unit presents an overview of recombinant protein purification with special emphasis on proteins expressed in *E. coli*. Practical aspects and strategies are stressed throughout, and wherever possible the discussion is cross-referenced to the example protocols described in the rest of Chapter 6.

The first section deals with information pertinent to protein purification that can be derived from translation of the cDNA sequence. This is followed by a brief discussion of some of the common problems associated with bacterial protein expression (see also *UNIT 5.1*). Planning a protein purification strategy requires that the solubility of the expression product be determined; it is also useful to establish the location of the protein in the cell—e.g., cytoplasmic or periplasmic. This unit includes flow charts that summarize approaches for establishing solubility and localization of bacterially produced proteins (see also *UNIT 5.2*).

Purification strategies for both soluble and insoluble proteins are reviewed and summarized in flow charts (see also Chapter 1). Many of the individual purification steps, especially those involving chromatography, are covered in detail in Chapters 8 and 9, and elsewhere (Janson and Rydén, 1989; Scopes, 1994). The methodologies and approaches described here are essentially suitable for laboratory-scale operations. Large-scale methodologies have been reviewed by Asenjo and Patrick (1990), Thatcher (1990a), and Walsh and Headon (1994a).

A section on glycoproteins produced in bacteria in the nonglycosylated state is included to emphasize that, although they may not be useful for in vivo studies, such proteins are well suited for structural studies. The final sections deal with protein handling, scale and aims of purification, and specialized equipment needed for recombinant protein purification and characterization.

PROTEIN SEQUENCE AND COMPOSITION ANALYSIS

Analyzing the Protein Sequence

The protein sequence translated from the DNA coding sequence is usually available, and before attempting any laboratory work it is useful to carry out a literature survey and basic computer analyses (Chapter 2). First, if the natural protein has been isolated and characterized, reviewing the protein's physicochemical properties and the purification techniques used may aid in planning a strategy for isolation from the recombinant host. For recombinant proteins that require folding from inclusion bodies (Bowden et al., 1991; Kane and Hartley, 1991; *UNIT 6.4*), information on the conformational properties, including denaturation/folding curves, can help rationalize folding scenarios.

Second, for uncharacterized proteins, analyses of related proteins with sequence similarities or known motifs may give useful clues for selecting purification steps (*UNIT 2.1*; see also the PROSITE database: Dictionary of Protein Sites and Patterns, Amos Bairoch, University of Geneva). For example, if the protein contains the well-known kringle domain, one may try lysine affinity chromatography (Cleary et al., 1989). On the other hand, if the protein contains no recognizable motifs and shows no similarity to other proteins, yet contains many cysteine residues, other strategies and precautions would be warranted as described in *UNITS 6.3-6.5*.

The amino acid sequence can be used to direct the synthesis of peptides corresponding to potential epitopes (e.g., 10 to 20 residues; *UNIT 2.2*). Polyclonal antibodies raised against the peptides may be suitable for detecting the protein of interest by immunoblotting. This approach may be especially valuable for monitoring proteins expressed at low levels—e.g., when *E. coli* secretion vectors are used. The antibodies may also be useful for immunoaffinity chromatography.

Analyzing the Amino Acid Composition

The amino acid composition (*UNIT 3.2*) of the protein will also allow calculation of some basic physicochemical parameters. First, using

Copyright © 1995 by John Wiley & Sons, Inc.

BCU LIBRARY OF CENTRAL LANCASHIRE UNIVERSITY

average pK_a values for ionizable side chains in proteins (Matthew et al., 1978), the isoelectric point (pI) can be estimated by applying the well-known Henderson-Hasselbach relationship. The calculations can be done with an electronic spreadsheet such as Lotus 123 or Excel (Atkinson et al., 1989) or with a commercially available software package (e.g., PC/Gene supplied by IntelliGenetics/Betagen). The values obtained, although only approximate, are useful for guiding the initial selection of ion-exchange resins and the pH of column buffers. When eukaryotic hosts are selected for protein expression, it should be noted that post-translational modifications such as phosphorylation and glycosylation will affect the pI.

Another parameter that can be estimated from the amino acid composition is the extinction coefficient (ε), usually at 280 nm (Wetlaufer, 1962). Although this information will be more useful when the protein has been purified, as most columns are monitored by UV absorption, proteins with an unusually low ε (no tryptophan and little or no tyrosine) may be difficult to detect during the early stages of purification.

Other physicochemical parameters that can be calculated include hydrodynamic parameters such as molecular radii and sedimentation coefficients (Laue et al., 1992). These parameters may help in interpreting results of gel-filtration and centrifugational separations.

CHARACTERISTICS OF THE HOST-VECTOR SYSTEM

Choosing an Expression System

Popular protein expression systems include *E. coli*, yeast, baculovirus-infected insect cells, and cultured mammalian cell lines (Chapter 5). If the requirement is to obtain a protein post-translationally modified via glycosylation (Chapter 12) or phosphorylation (Chapter 13), then a eukaryotic expression system must be used. Stable mammalian expression systems are the most time-consuming to establish and require the most expertise; however, they may be the only successful system for certain proteins, especially those that cannot be secreted in high enough yield, cannot be refolded from *E. coli*, or require the correct glycosylation pattern. The post-translational modifications may aid purification (e.g., lectin affinity chromatography can be used for glycoproteins; *UNIT 9.1*). On the other hand, these modifications may introduce charge heterogeneity—as is commonly observed with glycosylation (due to loss of sialic acid)—which may then complicate purification, especially with methods such as ion-exchange chromatography (*UNIT 8.2*). In this regard, purification from *E. coli* is the most straightforward, and most of the post-translational modifications observed in *E. coli* are nonspecific, such as deamidation (Wingfield et al., 1987a) and proteolytic clipping (Nagata et al., 1986). Other less common sources of protein heterogeneity arising from *E. coli* expression are (1) internal starts in translation (Dale et al., 1994), (2) partial readthroughs of the termination codon (Danley et al., 1991), and (3) translation errors (Lu et al., 1993).

The basic requirements for protein expression in *E. coli* are as follows:

1. Prepare the coding DNA (cDNA) sequence for the protein (i.e., without introns).
2. Insert the cDNA into an *E. coli* expression plasmid, thereby placing the gene under control of a promoter that is efficiently used by *E. coli* RNA polymerase (for examples of inducible promoters, see *UNITS 5.1 & 5.2*). The expression plasmid must also contain an efficient ribosome binding (translational initiation) site, a termination site (transcription terminator), and an antibiotic resistance gene (selectable marker). The plasmid and inserted gene are known as the expression vector.
3. Insert the expression vector into an appropriate *E. coli* host (transform cells).
4. Ferment and induce the transformed cells to produce the protein of interest (*UNIT 5.3*).

Considering the experience accumulated since the late 1970s, the expression of any given protein in *E. coli* is usually possible. Whether the protein will be expressed in a soluble and unmodified state, however, is unpredictable. The aforementioned elements of the expression system can be tinkered with to optimize protein production (*UNIT 5.1*).

Minimizing Proteolysis

If the protein is expressed in the cytoplasm in a soluble state, the purification can be carried out directly after cell lysis. Soluble recombinant proteins are, however, susceptible to proteolysis which can occur before or after extraction from the cell (Maurizi, 1992). Choosing protease-deficient *E. coli* host strains (Goff and Goldberg, 1985), manipulating growth conditions, especially the time of induction for inducible promoters (Allet et al., 1988), and using exogenous protease inhibitors can minimize this problem. Nevertheless, more extreme steps may be required, such as inducing the ex-

CU IAS CENTRAL UNIVERSITY OF KERALA

pressed protein to form insoluble inclusion bodies (Cheng et al., 1981; also see references cited in Kane and Hartley, 1991), using a secretion vector to locate the protein to the periplasm or medium, and changing to a eukaryotic expression system. In addition, there is a protein engineering approach, which requires knowledge of the proteolytic cleavage site(s): to stabilize the protein, one or both of the residues forming the scissile bond are changed by site-directed mutagenesis (Mildner et al., 1994).

Removing the Amino-Terminal Methionine

Another common problem with proteins expressed directly in *E. coli* is retention of the N-terminal methionine derived from the initiating *N*-formylmethionine (the formyl group is almost always removed). The N-terminal methionine is generally removed when the second amino acid is alanine, glycine, proline, serine, threonine, or valine (cleavable residues), but not when it is arginine, asparagine, aspartic acid, glutamic acid, glutamine, isoleucine, leucine, lysine, or methionine (noncleavable residues; Sherman et al., 1985). When recombinant proteins are expressed at very high levels, the N-terminal methionine can be retained regardless of the nature of the second amino acid, presumably due to saturation of the processing enzymes or depletion of required metal cofactors. Removal of cleavable N-terminal methionine can be carried out in vitro by digestion with purified methionine aminopeptidase (Miller et al., 1987) or by coexpression of the processing enzyme (Ben-Bassat et al., 1987).

The need to remove noncleavable methionines can be circumvented by incorporating an N-terminal secretion leader sequence that localizes the protein, minus the leader, to the periplasmic space (Holland et al., 1990). Other approaches utilize the incorporation of N-terminal fusions with, for example, ubiquitin, which can be cleaved in vitro or in vivo with the processing enzyme (ubiquitin hydrolase); this approach involves coexpression of the hydrolase (Miller et al., 1989). Finally, it should be noted that it is sometimes possible to resolve proteins containing N-terminal Met from those lacking it by chromatographic methods (Wingfield et al., 1987b).

Dealing with Inclusion Bodies

As previously mentioned, the expression of proteins in *E. coli* often results in the formation of inclusion bodies (*UNITS 6.3 & 6.5*). Inclusion bodies can be easily observed by phase-contrast microscopy as dense bodies, usually located at the polar extremities of the cell. The formation of inclusion bodies can occasionally protect proteins against proteolysis and can also allow accumulation of proteins normally toxic to the cell, such as human immunodeficiency virus type 1 (HIV-1) protease (Cheng et al., 1990). The formation of inclusion bodies also simplifies purification of the protein, albeit in a denatured/aggregated state (see below). The main disadvantage is that the protein must be extracted with protein denaturants and then folded into a nativelike conformation. For small (10- to 20-kDa) single-domain proteins, this is usually not problematical, although the overall recoveries may only be 5% to 20% those of similar or identical proteins expressed in a soluble state. For large (40- to 70-kDa), multiple-domain proteins, recoveries may be negligible, although there have been a number of successful cases, such as the 69-kDa tissue-plasminogen activator (Grunfeld et al., 1992).

The formation of inclusion bodies is not an intrinsic property of a particular protein and can sometimes be prevented. Approaches include the following: changing the promoter, host strain, and combinations thereof; controlling the growth conditions (especially the pH of the culture); adding nonmetabolizable sugars such as sucrose and sorbitol to the fermentation medium; and changing the temperature of induction, usually by lowering it (for reviews, see Schein, 1989; Wetzel, 1992).

The recombinant protein may be located in both the insoluble and soluble fractions of the cell (mixed-phase expression), and in these cases, better yields may be obtained by processing the soluble material (discarding the insoluble) even though it might constitute only a minor portion of the total expressed protein (Thatcher and Panayotatos, 1986; Wingfield et al., 1987c). Soluble protein purified from mixed-phase expressions should be carefully analyzed to check its authenticity (e.g., by mass spectrometry) as the solubility may have resulted from minor modifications such as deamidation or proteolysis of a few residues from either the N or C terminus (P.T.W., unpub. observ.).

The most successful approach for avoiding inclusion body formation is the use of an appropriate secretion vector. The N-terminal secretion signal directs protein to the periplasmic space (see section on Localizing Protein), and translocation across the plasma membrane results in cleavage of the secretion leader se-

BOULDER CENTRAL UNIVERSITY LIBRARY

quence. The environment of the periplasm contains enzymes that accelerate folding and the formation of disulfide bonds (Hayano et al., 1991; Missiakas et al., 1994; Zapun et al., 1995). Purification is also simplified as the protein content of the periplasmic space constitutes only 4% of the total *E. coli* protein (Beacham, 1979).

Engineering Fusion Proteins

Apart from direct expression, there are many examples of fusion protein expression. Fusion proteins consist of the protein of interest with N- or C-terminal extensions of known protein domains or short peptide sequences (*UNIT 5.1*; Uhlén et al., 1992). The fusion partner may help maintain the solubility of a protein which is normally expressed in insoluble form (LaVallie et al., 1993). The fusions may afford affinity handles for protein purification, can be useful carriers for short peptides (e.g., epitope regions) that are normally proteolytically degraded, and, as stated above, can be used to obtain authentic N-terminal sequences.

The expression of fusion proteins with affinity handles, such as those containing stretches of polyhistidine (His-tagged; see also *UNIT 6.5*), has become extremely popular due to the ease of protein purification under both nondenaturing and denaturing solvent conditions. The soluble fusion proteins often have native-like conformations and are biologically active. It cannot be assumed, however, that a tag will have no effect on the protein's function or activity. From a protein purification viewpoint, the main advantage of affinity-tagged proteins is realized when they are combined with a secretion vector (Skerra et al., 1991): in such a case the protein will be translocated to the periplasm or the medium, though often at low concentrations, and the tagged protein can be readily purified from culture medium or from osmotically shocked cells.

Table 6.1.1 briefly summarizes some of major advantages and disadvantages of the various expression scenarios using *E. coli*. If attempts to express protein in *E. coli* or to fold inclusion body protein fail, then a eukaryotic system must be considered. The decision of which system to use is often dictated by the expertise available to the laboratory. Alternatively, many companies offer custom expression services using various protein expression systems—this can be an expedient, although expensive, solution.

SOLUBILITY AND LOCATION OF THE PROTEIN

Determining Solubility

Figure 6.1.1 shows a simple centrifugation scheme that indicates how to determine the solubility of a protein expressed in *E. coli* (see also *UNIT 5.2*). The recombinant protein in the various fractions is assayed by SDS-PAGE (*UNIT 10.1*); if more sensitive methods are required, immunoblotting or biological assays may be used.

Cell breakage carried out with a French pressure cell (*UNIT 6.2*) will disrupt both the outer and inner membranes. The peptidoglycan layer, which lies underneath the outer membrane in gram-negative bacteria such as *E. coli*, will be fragmented into sheets. Low-speed centrifugation (30 min at 10,000 × *g*) separates unbroken cells, bacterial outer membrane and peptidoglycan components, and highly aggregated inclusion body proteins (pellet fraction) from soluble bacterial proteins, soluble recombinant proteins, and polymeric materials, including ribosomal protein complexes and inner membrane vesicles (supernatant fraction). High-speed centrifugation (90 min at 100,000 × *g*) of the low-speed supernatant will pellet polymers; soluble proteins, derived mainly from the cytoplasm and periplasmic space, can then be recovered from the clarified supernatant. Soluble proteins in the low-speed or high-speed supernatant are purified directly using conventional methods (*UNIT 6.2*).

It should be noted that a working definition of solubility is the presence of protein in the supernatant after centrifugation for 100 min at 100,000 × *g*. This definition applies to solvents of viscosity or density close to that of water.

Occasionally, recombinant protein will be found in both the pellet and supernatant fractions after low-speed centrifugation due to the accumulation of both soluble and inclusion body proteins. If partitioning is observed only following high-speed centrifugation, then specific self-association or nonspecific association involving *E. coli* proteins and nucleic acid may be suspected. Recombinant proteins that normally bind RNA or DNA often bind nonspecifically to bacterial nucleic acid (Sherman and Fyfe, 1990; Wingfield et al., 1990).

Inclusion body protein, which is located in the low-speed pellet fraction, can be partially purified by extracting with a mixture of detergent [usually 1% to 5% (v/v) Triton X-100] and denaturant, either urea or guanidine·HCl. The

PCU IASTRI UNIVERSITY LIBRARY

Table 6.1.1 Protein Expression in *Escherichia coli*

Protein expression[a]	Location[b]	Soluble	Advantages	Disadvantages
Native sequence	C	Yes	High-level expression Direct purification with good recovery	N-terminal Met may be retained Susceptible to proteolysis
	C	No	High-level expression May protect against proteolysis Toxicity effects of protein to cell may be avoided Easy partial purification (washed pellets—see Fig 6.1.1)	Protein folding must be carried out Recovery of purified native protein can be low to zero N-terminal Met may be retained
Native–fusion	C	Yes	High-level expression Purification aided with affinity-tagged protein Solubility and stability of expressed protein (or peptide) can be enhanced by fusion partners Authentic N terminus after site-specific cleavage (not always true)	To obtain native sequence site-specific cleavage of fusion protein required Overall yield of native protein may be low
	C	No	See comments for native insoluble Purification of denatured protein aided (e.g., His-tag) Protein folding may be enhanced by binding to affinity matrix (Sinha et al., 1994; *UNIT 9.4*)	See comments for native insoluble Protein folding required
Native–secretion	C	No		Secretion leader unprocessed, purification usually not attempted
	P/M	Yes	Ease of purification Correct N terminus Protein folded and oxidized	Expression level and recovery may be low
	P	No	Correct N terminus Expression level may be high May be protected against proteolysis	Protein folding must be carried out

[a]Native, native protein sequence, including any site-specific mutations or deletions; Native–fusion native sequence with N- or C-terminal extension sequence (e.g., polyhistidine tag); Native—secretion, native sequence plus N-terminal leader sequence coding for an *E. coli* secretory signal (e.g., OmpA).

[b]C, cytoplasm; P, periplasmic space; M, medium.

concentration of denaturant used for pellet washing is determined empirically and should be below the concentration required for solubilization of the recombinant protein; the usual ranges are ~1 to 4 M urea and ~ 0.5 to 1.5 M guanidine·HCl. The cloudy extract will consist of complex carbohydrate from the fragmented peptidoglycan layer, lipopolysaccharide, and outer membrane proteins. The inclusion body protein in the washed pellets is then extracted with solvents that disrupt protein-protein interactions (e.g., 6 to 8 M urea or guanidine·HCl) and processed as described below.

Localizing Protein

When proteins incorporating a secretion vector are expressed in *E. coli*, advantage can be taken of the fact that the recombinant proteins will be located in the periplasmic space and/or the culture medium. Secretion

BCS LIBRARY / CENTRAL UNIVERSITY LIBRARY

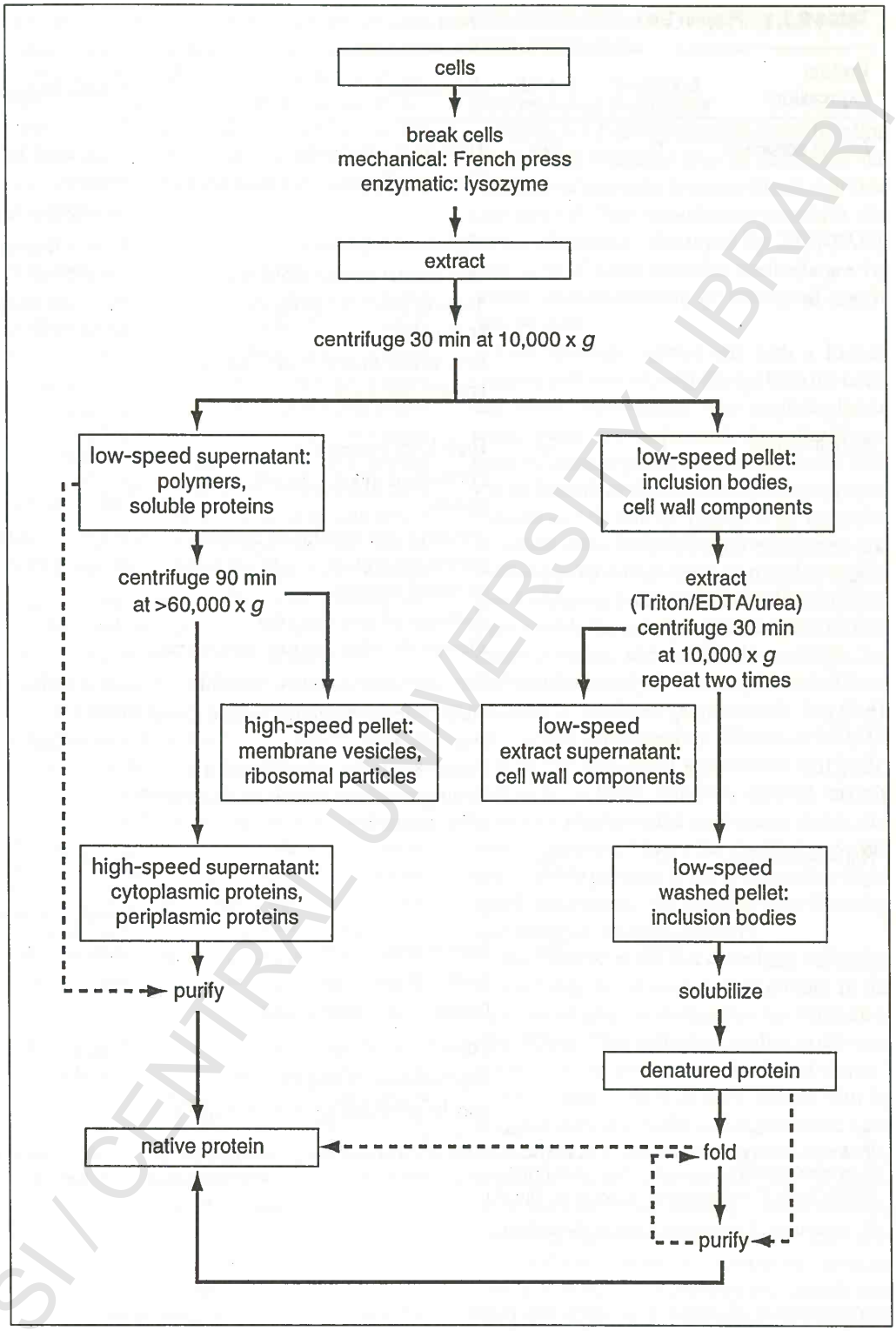


Figure 6.1.1 Differential centrifugation of *E. coli* cell lysates. Cells are broken with a French press or by lysozyme treatment. Insoluble (inclusion body) proteins, from either the cytoplasm or periplasm, are located in the low-speed pellet, which is subjected to preextraction to remove outer membrane and peptidoglycan material. Inclusion bodies are extracted from washed pellets with strong protein denaturants such as guanidine·HCl. The solubilized protein, which is denatured and reduced (free sulfhydryl residues), is either directly folded and oxidized (disulfide bonds formed) or purified before folding. Soluble proteins (from the periplasm and cytoplasm) are located in the low-speed and high-speed supernatants. The latter can be used directly for chromatography, whereas the former requires clarification by other techniques such as ammonium sulfate fractionation or membrane filtration.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

into the medium is due to "leakage" from the periplasm and appears to depend on the level of accumulation and the fermentation conditions. Figure 6.1.2 summarizes approaches used to recover proteins selectively from the periplasmic space or the medium (see also *UNIT 5.2*). High-level secretion into the periplasm sometimes results in the formation of aggregates, analogous to cytoplasmic inclusion bodies (Bowden et al., 1991). Periplasmic inclusion bodies can be extracted from the low-speed pellet fraction following normal cell breakage (see Fig. 6.1.1).

Proteins in the medium can be recovered by subjecting the culture medium to centrifugation or filtration, steps that remove intact cells and any large debris. The clarified protein is usually dilute and is often concentrated prior to purification by affinity or conventional chromatography. Periplasmic proteins can be selectively released by osmotic shock (preferred method) or by selective disruption of the outer membrane and peptidoglycan layer using lysozyme.

Apart from its use in dissecting the bacterial compartments, lysozyme is often employed to prepare complete cell lysates, especially in laboratories that do not have access to a French pressure cell. Cells treated with lysozyme can be disrupted with detergents or by brief sonication (*UNIT 6.5*).

Useful microscale (<1 ml) *E. coli* cell fractionation schemes have been based on osmotic shock treatment (Yarranton and Mountain, 1992) or repeated freezing and thawing of cells (Johnson and Hecht, 1994). *UNIT 5.2* gives small-scale (1 to 25 ml) procedures for preparing samples of periplasmic extracts and extracellular media for analysis by SDS-PAGE (*UNIT 10.1*).

STRATEGIES FOR ISOLATION OF SOLUBLE PROTEINS

There are no set formulas for isolating soluble recombinant or nonrecombinant proteins; there are, however, some basic strategies and precautions that can be followed. A flow chart summarizing some of the methods commonly used for *E. coli* is shown in Figure 6.1.3 (see also Chapter 1), and a specific protocol detailing the purification of interleukin 1β (IL-1β) is presented in *UNIT 6.2*. Comments on the various stages are given in order of their application.

Determining the Isoelectric Point

The section on Analyzing the Amino Acid Composition mentions how the isoelectric point of a protein can be estimated from the pK_a values for ionizable side-chain groups. The pI can also be determined experimentally by subjecting the soluble protein extract to one-dimensional isoelectric focusing (*UNIT 10.2*) or two-dimensional titration curve analysis (Watanabe et al., 1994). If the recombinant protein is not a major protein in the extract, specific detection by immunoblotting will be required. The calculated pI can be used to set the buffer pH in subsequent ion-exchange steps.

Breaking Cells

Cells are efficiently broken by high-pressure homogenization using a continuous-fill French pressure cell, which is suitable for processing volumes of 40 to 250 ml (reviewed by Hopkins, 1991; see *UNIT 6.2*). For volumes exceeding 500 ml, the Manton-Gaulin-APV homogenizer (APV Gaulin) should be used. Sonication is also useful for breaking cells but is best suited for volumes below 100 ml. Alternatively, the outer cell wall can be enzymatically digested with lysozyme (200 μg/ml) and the cells broken by detergents, sonication, or both (Kaback, 1971; Burgess and Jendrisak, 1975; *UNIT 6.5*). Proteins that are secreted into the periplasmic space can be selectively released by hypotonic (osmotic) shock (Heppel, 1967).

The viscosity of the cell lysate may be high due to released nucleic acid. Before centrifugation, the viscosity must be reduced either by sonicating or by adding DNase (25 to 50 μg/ml plus 5 to 10 mM Mg^{2+}) and RNase (50 μg/ml; no Mg^{2+} requirement). A standard protease inhibitor mixture should be included in the buffer—containing, for example, 2 to 5 mM EDTA, 0.5 to 1.0 mM phenylmethylsulfonyl fluoride (PMSF) or 5 mM benzamidine, and 1 μM pepstatin A. The serine protease inhibitor 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF) is a water-soluble substitute for PMSF with a much longer half-life in aqueous solution and is used at ~50 μM. (Boehringer Mannheim publishes a useful brochure entitled "Protease Inhibitors" that lists properties and handling conditions for most commercially available reagents). The addition of α_2-macroglobulin (1 μg/mg recombinant protein) before the final purification step(s) can protect protease-sensitive proteins (Ultsch et al., 1991). The crude extracts should be kept cold and the recombinant protein taken rapidly to a stage of the purification process where it is stable against contaminating proteases (e.g., as an ammonium sulfate precipitate).

BOSTON CENTRAL UNIVERSITY LIBRARY

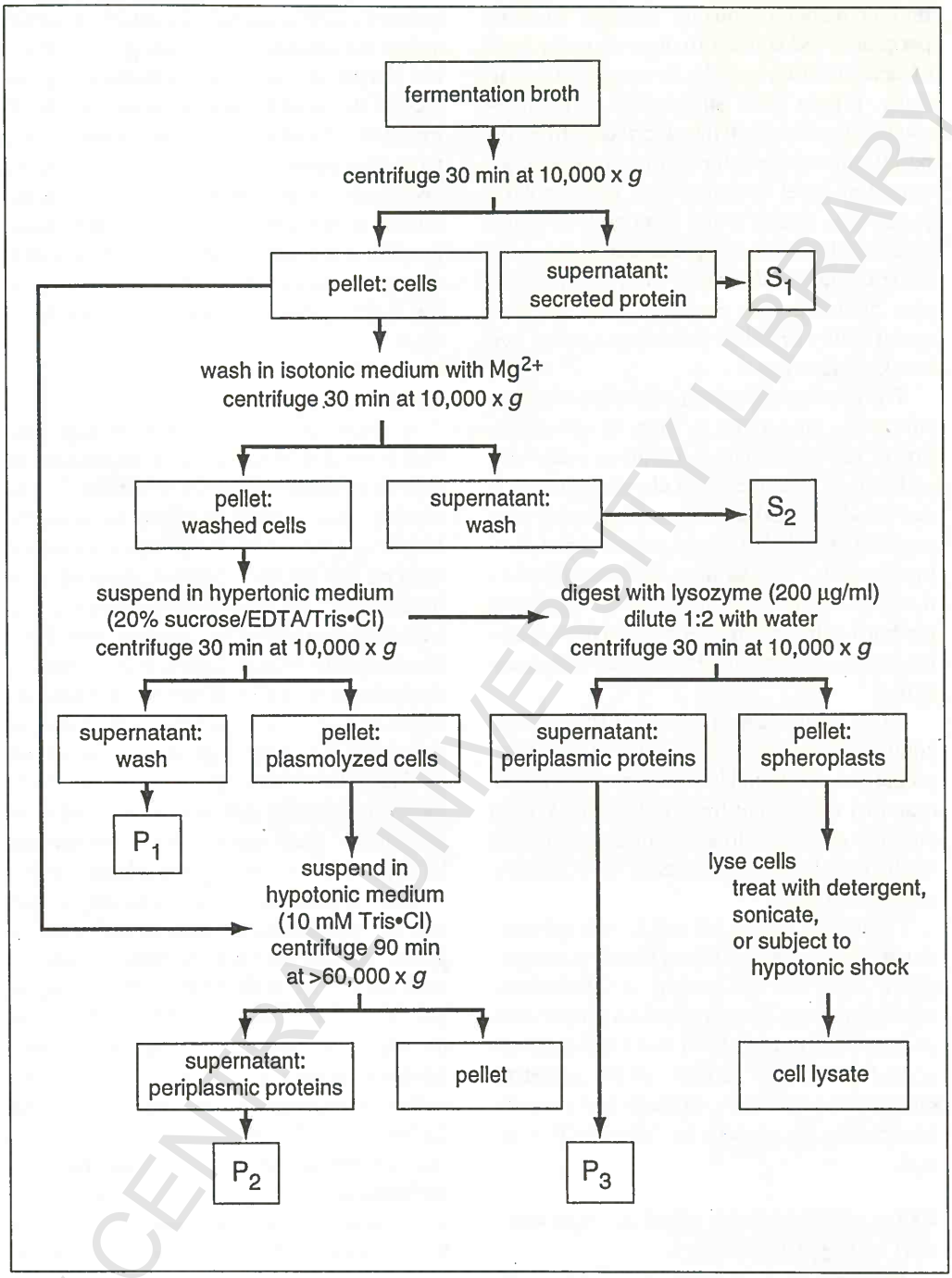

fermentation broth
centrifuge 30 min at 10,000 x g
pellet: cells
supernatant: secreted protein
S1
wash in isotonic medium with Mg2+
centrifuge 30 min at 10,000 x g
pellet: washed cells
supernatant: wash
S2
suspend in hypertonic medium (20% sucrose/EDTA/Tris•Cl) centrifuge 30 min at 10,000 x g
digest with lysozyme (200 µg/ml) dilute 1:2 with water centrifuge 30 min at 10,000 x g
supernatant: wash
pellet: plasmolyzed cells
supernatant: periplasmic proteins
pellet: spheroplasts
P1
suspend in hypotonic medium (10 mM Tris•Cl) centrifuge 90 min at >60,000 x g
lyse cells treat with detergent, sonicate, or subject to hypotonic shock
supernatant: periplasmic proteins
pellet
cell lysate
P2
P3

Figure 6.1.2 Localization of secreted and periplasmic proteins in *E. coli.* Periplasmic protein produced via a secretion vector can leak into the medium and be recovered by centrifugation (supernatant, S_1) or filtration. Washing cells with an isotonic solution such as lightly buffered 0.15 M NaCl or 0.25 M sucrose can also release protein (S_2). The compartmentalized periplasmic proteins are released by isotonic shock treatment by directly suspending normal cell paste or plasmolyzed cell paste into hypotonic medium. Plasmolyzed cell paste is derived by suspending cells in hypertonic medium and then pelleting. (In hypertonic medium the cell contracts, separating the inner membrane from the cell wall, and is said to be osmotically sensitized.) The hypertonic wash often releases protein (P_1). The supernatant from shocked cells (P_2) will contain constitutive *E. coli* proteins and the recombinant product. Osmotically sensitized cells can also be treated with lysozyme to fragment the outer membrane, thus releasing periplasmic proteins (P_3). The pellet from the lysozyme treatment contains spheroplasts (cells with fragmented outer membranes), which are easily disrupted by detergents, sonication, or hypotonic shock to release cytoplasmic proteins.

ECU IAS CENTRAL UNIVERSITY LIBRARY

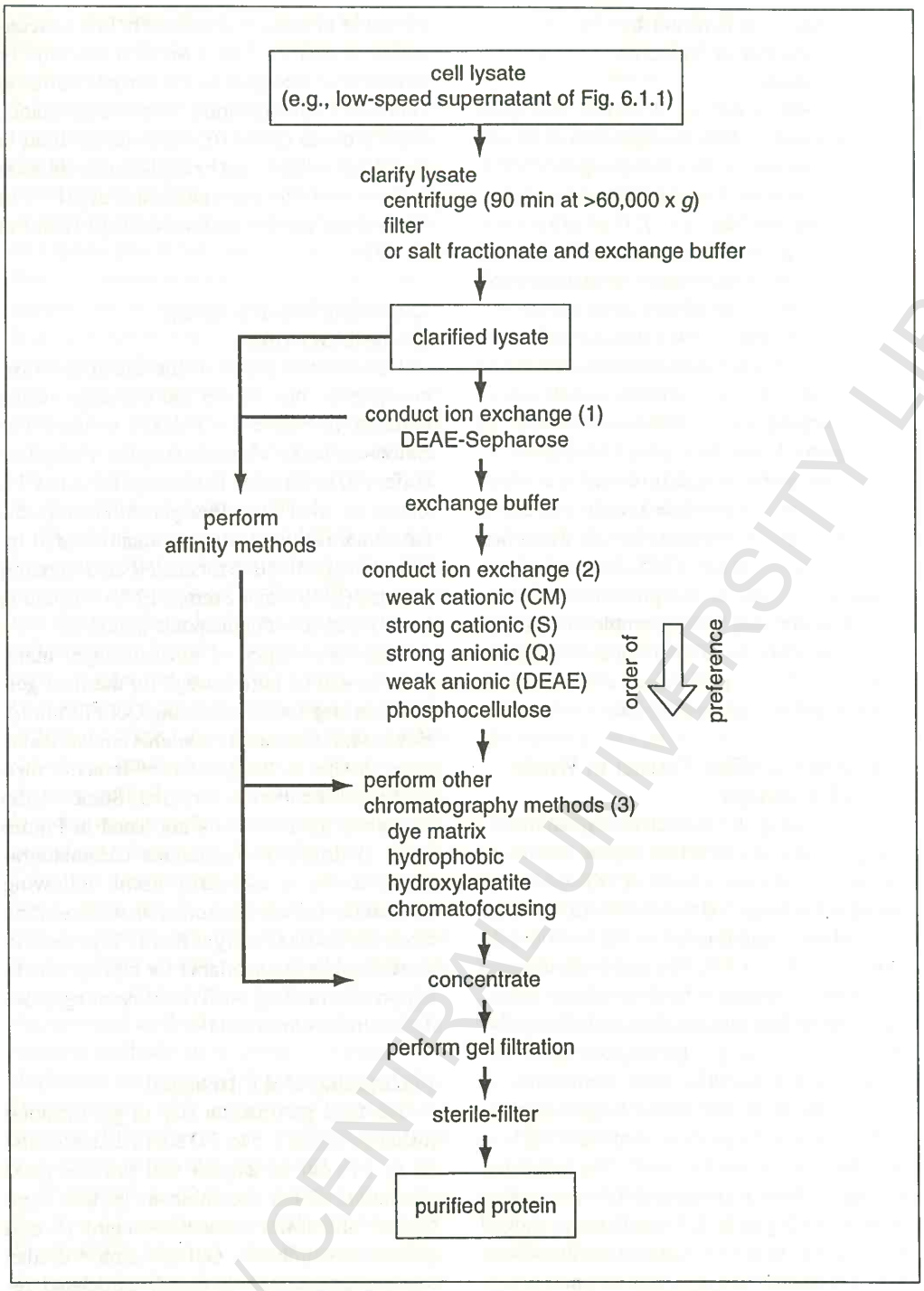


Figure 6.1.3 Purification of soluble proteins from *E. coli* lysates. Abbreviations for ion-exchange resins are as follows: CM, carboxymethyl; DEAE, diethylaminoethyl; Q, quaternary ammonium; S, methyl sulfonate. The order of preference for the stages of ion-exchange (2) and other methods (3) is based on the author's opinion and does not necessarily represent a consensus view. On the other hand, the use of a DEAE-based matrix at an early stage (1) is common practice. Affinity methods (see text and Chapter 9) can be performed at any stage following clarification of the lysate.

B. C. U. „M. ... IAŞI
C H

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

Clarifying Cell Extract by Centrifugation or Selective Precipitation

The lysate is subjected first to low-speed centrifugation to remove unbroken cells and large cellular debris, then to high-speed centrifugation to remove ribosomal material and other particulates (see Fig. 6.1.1). If an ultracentrifuge is unavailable, the extract can be clarified by the following techniques: ammonium sulfate or polyethylene glycol fractionation (reviewed by Scopes, 1994), phase partitioning (reviewed by Walter and Johansson, 1986), and membrane filtration (Walsh and Headon, 1994a; useful guides on filtration methodology are available from Millipore and Amicon).

Proteins that bind tightly to nucleic acid can be selectively precipitated with polyethyleneimine and resolubilized by salt extraction (Burgess and Jendrisak, 1975). In practice, particular properties of the protein can be exploited at this stage; for example, the protein of interest may be soluble under conditions where most *E. coli* proteins are insoluble, such as acidic pH or high temperature.

Applying Clarified Extract to Weak Anion Exchanger

Fractionating the extract with an anion-exchange resin is a useful first step as it removes host *E. coli* proteins, many of which have pI values in the range 5.0 to 7.0 and will thus bind to a column equilibrated in 50 to 100 mM Tris·Cl, pH 7.5 to 8.0. The positively charged matrix will also tightly bind nonproteinaceous materials such as nucleic acids and other polyanionic species (e.g., lipopolysaccharide derived from the bacterial outer membrane). A useful cleanup of the protein will take place whether or not the protein of interest binds to the column (see *UNITS 6.2 & 6.5*). The following column sizes are recommended for processing extracts: for 5 g cells, 2.5-cm diameter packed to a height of 10 to 15 cm; for 50 g cells, 5.0-cm diameter packed to a height of 20 cm.

Preparing for Repeat Ion-Exchange Step

Before repeating ion exchange, the solvent pH and ionic strength usually need adjustment. This can be carried out by dialysis (*UNIT 6.2*) or by gel filtration on a desalting column using, for example, Sephadex G-25 or G-50 (*UNIT 8.3*). In preparation for cation-exchange chromatography, dialysis against slightly acidic buffers (pH 5.0 to 6.0) will result in the helpful precipitation of some *E. coli* proteins. It may also be advisable to include a relatively low concentration of urea (0.5 to 2 M) or a nonionic or zwitterionic detergent in the dialysis buffer to minimize coprecipitation with contaminants. Basic proteins (pI >9.0), which do not bind to the DEAE column, can be applied after dilution to a cation exchanger equilibrated at pH 7.0 to 7.5 without careful buffer exchange (Allet et al., 1988).

Repeating Ion-Exchange Chromatography

For a second round of ion-exchange chromatography, one of the ion-exchange resins indicated in Figure 6.1.3 should be used. For cation-exchange chromatography, phosphate buffer (10 to 50 mM) between pH 5.0 and 7.5 should be tried first. Phosphocellulose (a bifunctional cation exchanger manufactured by Whatman) is effective for nucleic acid–binding proteins (Kelley and Stump, 1979). Protein is usually eluted with phosphate gradients.

After two stages of ion exchange, many proteins will be pure enough for the final gel-filtration step (see Performing Gel Filtration). However, if the sample contains contaminants close in size to the protein of interest, then further purification is required. Some of the frequently used methods are listed in Figure 6.1.3. Hydrophobic-interaction chromatography (*UNIT 8.4*) is especially useful following ammonium sulfate fractionation or salt elution from an ion-exchange resin. Dye-matrix screening kits are available for rapidly checking protein binding on five different agarose-dye matrices (Amicon).

Performing Gel Filtration

The final purification step of gel filtration (using a column 1.5 to 5.0 cm in diameter and 60 to 100 cm in length) will provide good separation of the recombinant protein from higher- and lower-molecular-weight *E. coli* protein contaminants. Gel filtration will also separate aggregated or highly associated recombinant protein from the physically stable form of the protein (e.g., monomer or dimer). Finally, it will be possible to exchange the buffer. The protein solution is usually concentrated before being applied to the column. After chromatography, the protein will be diluted 3- to 5-fold (or more) and may therefore require repeat concentration.

Other Methods

In addition to the generalized approach described, affinity methods can be applied at any

ECU IASL CENTRAL UNIVERSITY LIBRARY

stage following clarification of the extract. Biospecific affinity can be exploited with immobilized natural ligands such as antibodies, substrates, and receptor ligands. Affinity chromatography, which selects for particular classes of proteins, is carried out with immobilized lectins (for glycoproteins), dyes (for nucleotide-binding proteins) and nucleic acid or heparin (for RNA- and DNA-binding proteins). Commercially available antibodies against post-translationally modified residues (e.g., phosphotyrosine) are also useful. The application of affinity tags or fusions has been previously mentioned. Affinity methods are most useful when high degrees of purification are required—e.g., for proteins secreted into the medium, for small-scale isolations, or for rapid purification requirements.

The most commonly used affinity method is immunoaffinity chromatography. The ideal reagent is a monoclonal antibody that has been specifically selected to have a moderate-to-low affinity for the ligand in question, thus allowing elution under nondenaturing conditions. Antibodies raised against peptides often have lower affinities for the native protein than antibodies raised against the intact protein. Elution from peptide-antibody immunoaffinity columns can be achieved using the competing immunizing peptide (reviewed by Sutcliffe et al., 1983). Directed immobilization of the antibody, where only the Fc domain is bound to the column matrix and the antigen binding site (Fab domain) is thus oriented away from the matrix, results in higher binding efficiencies. An oriented antibody matrix can be made by binding antibody to immobilized protein A–Sepharose (or protein G–Sepharose) and fixing it in position with a covalent cross-linking reagent (Schneider et al., 1982; commercial kits can be obtained from Pierce).

Compilations of standard chromatographic fractionation media are available (Table 8.2.2; Patel, 1993). However, the best source of information is often the literature from the various manufacturers.

STRATEGIES FOR ISOLATION OF INSOLUBLE PROTEINS

Recombinant proteins expressed in *E. coli* that are located in the low-speed pellet fraction (see Fig. 6.1.1) following cell lysis are highly aggregated (i.e., inclusion bodies). Inclusion bodies are normally derived from protein aggregation in the cytoplasm, or in the periplasm if a secretion vector was used. As mentioned above, protein can also be located in either the low- or high-speed pellet fractions because of interaction with bacterial nucleic acids. Furthermore, if the protein is known to undergo polymerization in vitro (e.g., viral nucleocapsid subunits), expression in *E. coli* can also be expected to lead to polymerization in vivo to varying degrees, and such proteins will be partitioned in both the supernatant and pellet fractions (Wingfield et al., 1995). There are also examples of membrane proteins which, when expressed in *E. coli*, associate with the inner cytoplasmic membrane and can be extracted with nondenaturing detergents (Bibi and Béjà, 1994, and references cited therein).

When insolubility is due to interactions involving folded protein as described above, extraction under nondenaturing conditions should be attempted, for example, using salt (1 M NaCl) and nondenaturing detergents (10 mM CHAPS or 2% Triton X-100). Insolubility due to classic inclusion body formation requires extraction with denaturing solvents, and the remainder of this section deals with this subject. The flow chart in Figure 6.1.4 illustrates some of the approaches possible for processing protein extracted from inclusion bodies.

Breaking Cells

Cells can be broken by mechanical means (*UNIT 6.2*), enzymatically with lysozyme (*UNIT 6.5*), or by a combination of methods (*UNIT 6.5*). It is advantageous to break cells as completely as possible, as any unbroken cells will be located in the low-speed pellet fraction from which the recombinant protein will be extracted.

Preparing Washed Pellets

The object of the initial low-speed centrifugation is to extract as many *E. coli* contaminants as possible without solubilizing the recombinant protein. This is usually carried out as described in the section on Determining Solubility (see also Fig. 6.1.1).

When a fixed-angle rotor is used, pellets from the low-speed centrifugation consist of at least two light-colored layers and a darker, hard-packed pellet at the bottom of the tube (Fig. 6.1.5). The hard-packed material is probably a small amount of unbroken cells. The next layer is inclusion body protein, and the top layer (least dense and lightest in color) is outer membrane and peptidoglycan fragments. Analysis of the top layer by SDS-PAGE (after heating proteins in SDS sample buffer at >80°C) will reveal two strong bands at ~35 and 38 kDa

BOSTON UNIVERSITY CENTRAL LIBRARY

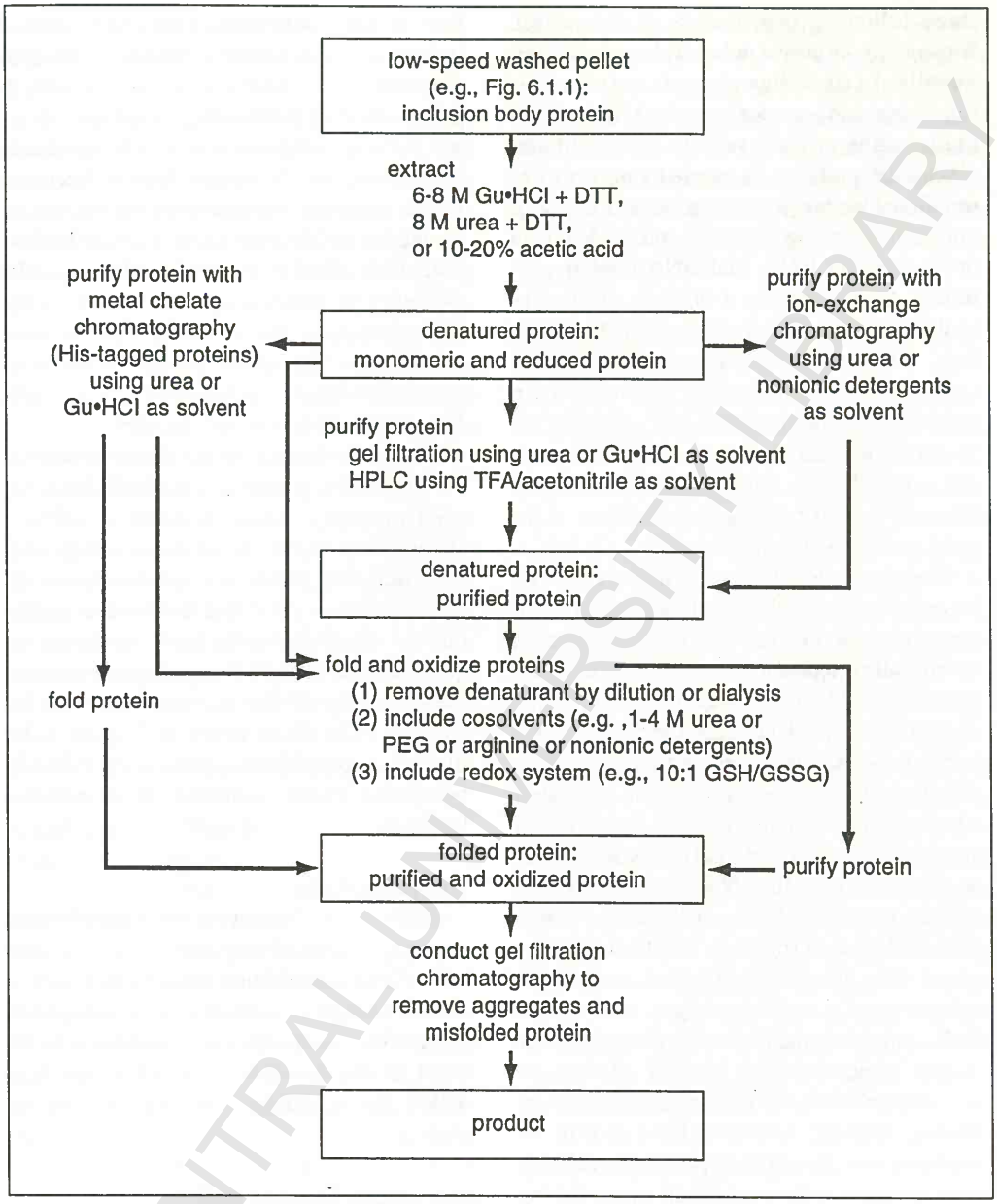


Figure 6.1.4 Folding and purification of inclusion body proteins from *E. coli*. The protein is extracted with protein denaturants such as guanidine·HCl (Gu·HCl), urea, or an organic acid. The reductant dithiothreitol (DTT) is included to prevent artificial disulfide bond formation (especially intermolecular bonds). The denatured protein can be purified by various methods and then folded, or it can be directly folded. Typically, some purification (e.g., gel filtration in Gu·HCl) prior to folding is recommended as it often results in higher folding yields. Protein folding and oxidation are carried out concurrently. Disulfide bond formation is catalyzed by low-molecular-weight thiol/disulfide pairs such as reduced (GSH) and oxidized (GSSG) glutathione. GSH/GSSG ratios of 5:1 to 10:1 are normally used, which are similar to those found in vivo in the endoplasmic reticulum (Hwang et al., 1992). A cosolvent is included to maintain solubility during folding. Folded protein is purified if necessary (purification is usually needed if the protein is directly folded). Gel filtration is a useful final step for removing aggregated and or misfolded protein.

representing OmpA and the matrix proteins OmpC and OmpF, respectively, from the outer membrane (DiRienzo et al., 1978; see also Fig. 6.3.1). The outer membrane/peptidoglycan layer can be partially removed by resuspending and centrifuging at reduced speed (or time) or by diluting the suspension. Alternately, the cells can be pretreated with lysozyme prior to the French press cell breakage as described in *UNIT 6.5*. The lysozyme treatment reduces the size of the loosely pelleted outer membrane/peptidoglycan material so it locates pre-

ECU IASI CENTRAL UNIVERSITY LIBRARY

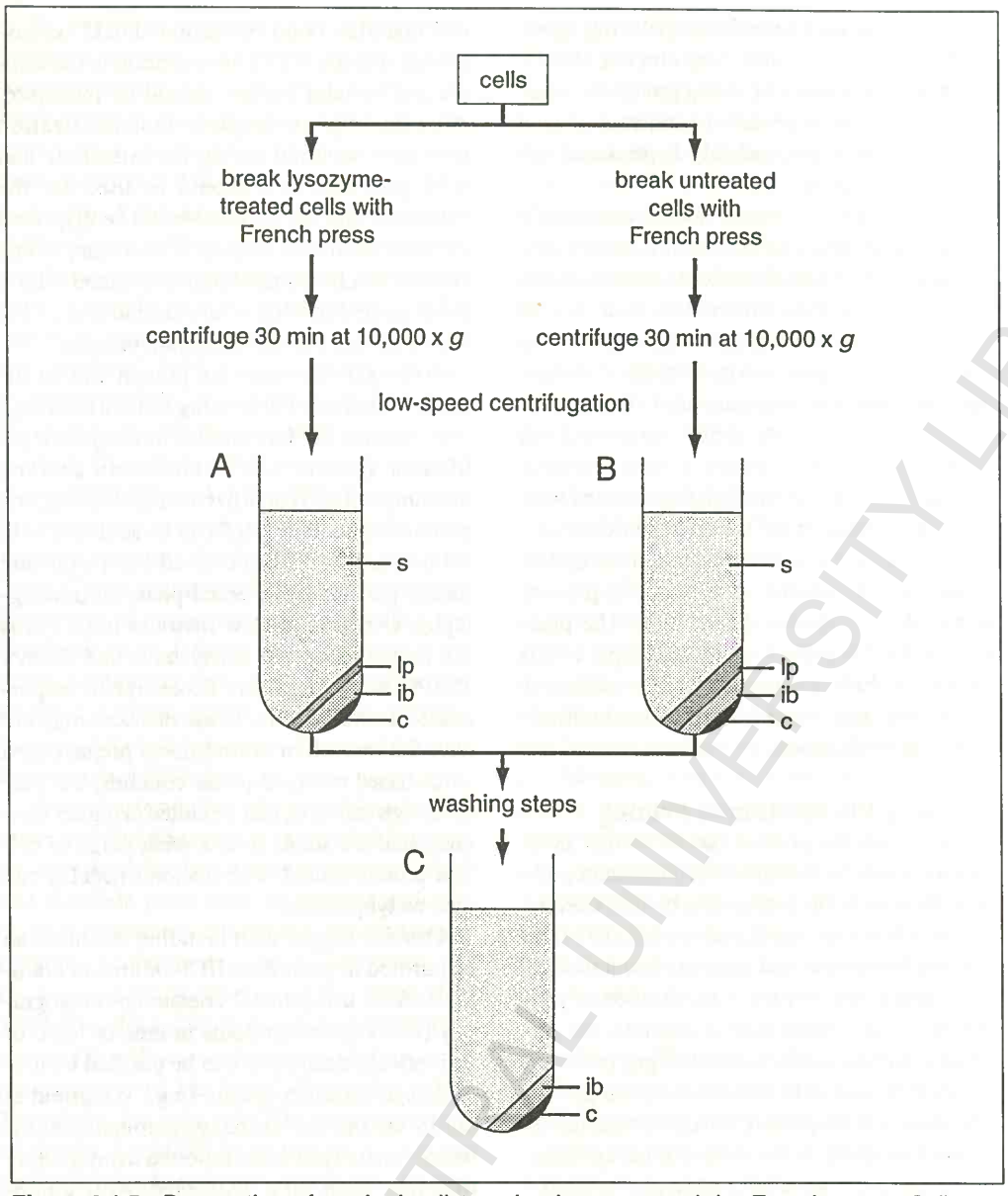


Figure 6.1.5 Preparation of washed pellets using lysozyme and the French press. Cells are broken with the French press with or without prior treatment with lysozyme. After low-speed centrifugation using a fixed-angle rotor, the contents of the centrifuge tubes have the characteristics shown. The contents of tubes A and B are labeled: s, supernatant; lp, loose pellet; ib, inclusion body protein; and c, unbroken cells and large cellular debris. The loose pellet material is derived from the outer cell wall and outer membrane (see text for further details). After washing the insoluble material (*UNIT 6.3*), the pellet should consist mainly of the inclusion body layer (tube C), and the supernatant should be fairly clear.

dominately in the low-speed supernatant (Fig. 6.1.5). The recombinant protein in a well-prepared washed pellet will be >60% pure when analyzed by SDS-PAGE (*UNIT 6.3*).

Extracting Protein

The washed pellets are extracted with high concentrations of protein denaturants such as 6 to 8 M guanidine·HCl or urea. It should be noted that some proteins are resistant to denaturation with high concentrations of these reagents, especially urea. Solubilization with 10% to 20% acetic acid has also been useful (*UNIT 6.5*); other denaturants for extracting inclusion bodies are described elsewhere (*UNIT 6.3*; Marston and Hartley, 1990). For background information on the suggested mode of action of protein denaturants, readers should consult the reviews of Tanford (1968) and Creighton (1993). If the protein contains cysteine resi-

BOU TESI CENTRAL UNIVERSITY LIBRARY

dues, it is essential to include a reducing agent, preferably 5 to 10 mM dithiothreitol (DTT). Even in the presence of strong protein denaturants, it may be necessary to sonicate or heat samples briefly to completely disperse and solubilize the protein.

The extraction process should completely unfold the protein and convert it to monomers. If a native protein can be unfolded, for example, with 4 M urea, the same protein in an *E. coli* inclusion body will almost certainly not be completely extracted in a monomeric, unfolded state with that same concentration of urea. The initial extraction trials should be carried out with guanidine·HCl, which is more effective than urea. Most proteins will be extracted with 6 to 8 M guanidine·HCl. There should be adequate reductant present to maintain sulfhydryl groups in the reduced state, and thus prevent artificial disulfide bond formation. The presence of EDTA and a slightly acidic pH of 6.0 to 6.5 will help minimize cysteine oxidation. The extract may require clarification by filtration or centrifugation.

Choosing Purification or Folding

The extracted protein can be further purified, or it can be directly folded and then purified. Protein folding may (Darby and Creighton, 1990) or may not (London et al., 1974) be affected by protein and nonproteinaceous contaminants in the extract. If purification of protein in the denatured state is possible, use purified material to develop a folding protocol. Apply the protocol to directly extracted protein and observe if the presence of contaminants has an adverse effect on the yield of folded protein. If not, directly extracted protein can be used.

Finally, there may be specific reasons for purifying proteins in the denatured state. For example, some proteolytic enzymes, such as HIV-1 protease, self-digest (undergo autoproteolysis) in the uninhibited state (Mildner et al., 1994, and references cited therein) but can be purified intact in the denatured (inactive) state, then refolded when required. It should be noted that, in general, unfolded proteins are more susceptible to chemical and proteolytic modifications.

Purifying Denatured Protein

If the protein is extracted with guanidine·HCl, gel filtration is a useful purification method; often protein >80% pure can be obtained (*UNIT 6.3*). The order of elution from the column should be directly proportional to the apparent molecular weight. If not, then artificial disulfide bond formation should be suspected, and the DTT concentration in the sample and column buffers should be increased. Alternatively, incomplete monomerization may have occurred during the extraction, and 8 M guanidine·HCl should be used for the extraction and the protein should be dispersed by sonication and heated if necessary). The column can be equilibrated and eluted with a lower guanidine·HCl concentration (e.g., 4 M) than that used for the extraction process.

After gel filtration, the protein can be directly folded (see Performing Protein Folding), or denaturant can be removed by dialysis or gel filtration against 1% to 5% (v/v) acetic acid and the sample then lyophilized. Alternatively, the protein in guanidine·HCl can be acidified with trifluoroacetic acid [TFA; ≤0.1% (v/v)] and further purified by reversed-phase chromatography. Useful high-flow matrices (e.g., Poros R2 from PerSeptive Biosystems and Source 15RPC from Pharmacia Biotech) can be purchased as bulk media. These matrices may not have the resolution of traditional prepackaged silica-based reversed-phase columns, but they have high capacity, can be eluted at higher flow rates, and are stable over a wide range of pH. The protein eluted with acetonitrile/TFA can then be lyophilized.

Proteins tagged with histidine residues can be purified in guanidine·HCl- or urea-containing buffers using metal chelate chromatography (*UNIT 6.5*), and proteins in urea or ionic or zwitterionic detergents can be purified by ion-exchange chromatography (e.g., Wingfield et al., 1990). For ion-exchange chromatography, better results have been reported using protein that has been first extracted with guanidine·HCl, then exchanged into urea (Shire et al., 1984).

If urea is used either for extraction or for maintaining solubility during refolding, a cyanate scavenger such as a glycine- or Tris-based buffer should be included to prevent carbamylation of the protein (Stark et al., 1960). For critical work, urea can be deionized with a mixed bed ion-exchange resin (see Protein Folding Reagents, *APPENDIX 3A*).

Performing Protein Folding

Protocols for folding proteins basically involve controlled removal of the denaturant under conditions that minimize aggregation and allow correct formation of disulfide bonds. For overviews of the practical aspects of protein folding, see *UNIT 6.4*, Jaenicke and Rudolph (1989), Thatcher (1990b), and Wetzel (1992).

ECU IASTE LIBRARY

To minimize nonproductive aggregation, folding is normally carried out at low protein concentrations (e.g., 0.01 to 0.10 mg/ml); for small, single-domain proteins, higher concentrations (e.g., 0.1 to 1.0 mg/ml) can often be used. Dilution and dialysis are the most common methods for removing the denaturant. Solubility during folding can be maintained with cosolvents such as nondenaturing concentrations of urea (1 to 4 M; London et al., 1974; *UNIT 6.5*) or guanidine·HCl (0.1 to 1.5 M; Orsini and Goldberg, 1978), arginine (Jaenicke and Rudolph, 1989), nonionic detergents and lipids (Zardeneta and Horowitz, 1994), cationic detergents (Puri et al., 1992), and polyethylene glycol (PEG; Cleland et al., 1992). Presumably, the cosolvents minimize intermolecular associations between the hydrophobic and thus "sticky" folding intermediates; for a theoretical discussion of aggregation versus folding, see Goldberg et al. (1991) and Kiefhaber et al. (1991). Additives such as ammonium sulfate, glycerol, sucrose, enzyme substrates or inhibitors, and ligands have also been used to enhance protein folding.

Oxidative formation of disulfide bonds (one of the rate-limiting steps in protein folding) can be catalyzed by low-molecular-weight thiol and disulfide pairs such as reduced and oxidized glutathione (GSH/GSSG). Redox buffers facilitate oxidation through thiol/disulfide exchange reactions (reviewed by Wetlaufer, 1984; Creighton, 1984). Normally GSH/GSSG ratios of 5 to 10 are used with a total glutathione concentration of 1 to 5 mM (Wetlaufer, 1984). To reduce the rate of GSH loss due to air oxidation, 1 mM EDTA should be included in the buffer (Wetlaufer et al., 1987). The optimal concentrations and ratios of reagents are established in an empirical manner. Folding and oxidation are normally carried out concurrently.

For examples of preparative protein folding, see *UNIT 6.5*. The refolding of Fab fragments expressed in *E. coli* (Buchner and Rudolph, 1992) is illustrative of the systematic and empirical approach used to optimize folding conditions. Other examples of interest are described by Kohno et al. (1990) and Grunfeld et al. (1992).

The in vivo folding and assembly of proteins are assisted by accessory enzymes such as protein disulfide isomerase, which catalyzes disulfide bond formation; peptidyl prolyl *cis-trans* isomerase, which catalyzes the rate-limiting rotation of X-Pro bonds during protein folding; and various molecular chaperones which prevent aggregation of folding intermediates (*UNIT 6.4*; reviewed by Jaenicke, 1993; Ellis, 1994). Amrein et al. (1995) produced an *E. coli* expression system containing two plasmids, one for overproduction of the bacterial chaperones GroEL and GroES and the other for expression of the recombinant protein, a tyrosine kinase. Coexpression of the chaperones and the kinase resulted in increased solubility of the latter (compared to when expressed alone). This approach may have general applicability for producing soluble (or partially soluble) forms of proteins that are otherwise produced in an insoluble state.

Purifying Folded Protein

Once the protein has been folded, any of the purification methods discussed in Chapters 8 and 9 can be used. The number of purification steps required should be fewer than those for a protein expressed in a soluble state because of the purification factor obtained by preparation of washed inclusion bodies (*UNIT 6.3*). One of the purification methods that should be included is gel filtration, which may be the only one required. A correctly selected matrix should remove any remaining *E. coli* proteins and separate aggregated and misfolded protein from the native folded protein. Misfolded protein may be expected to have a larger molecular radius (higher apparent mass) than the corresponding native protein.

Monitoring Protein Folding

The restoration of function (e.g., enzymatic or biological activity) is perhaps the best criterion for detecting successful folding. However, it is not always practical to use activity measurements to monitor folding. It is also worth mentioning that an unfolded protein may become activated following the dilution required for many activity measurements. Conversely, native proteins can be denatured or inactivated during prolonged incubation at 37°C or by adsorption to microtiter plates. The use of antibodies to monitor protein folding is briefly reviewed by Goldberg (1991), and a concise review of common spectroscopic methods, such as circular dichroism and fluorescence, is provided by Schmid (1989).

BACTERIAL EXPRESSION OF PROTEINS NORMALLY GLYCOSYLATED

Because *E. coli* lacks glycosylation machinery, expression of glycoproteins in *E. coli* systems results in the synthesis of nonglyco-

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

sylated variants. Glycoproteins expressed directly in *E. coli* are often, but not always, insoluble. In vitro folding studies with glycosylated and nonglycosylated forms of proteins indicate that the carbohydrate can stabilize folding intermediates, and thus enhance folding, while not necessarily affecting the stability of the native state (Kern et al., 1993, and references cited therein). In eukaryotic cells, interference with protein glycosylation can lead to the formation of misfolded, aggregated, and degraded protein. This indicates that in vivo glycosylation (N-linked) may also prevent the aggregation of folding intermediates (reviewed by Helenius, 1994).

Despite potential pitfalls, many nonglycosylated protein variants have been successfully folded from *E. coli* inclusion bodies. Examples include cytokines of biomedical importance such as granulocyte/macrophage colony-stimulating factor (GM-CSF; Diederichs et al., 1991) and interleukin 5 (IL-5; Milburn et al., 1993). Inclusion body formation was avoided in some studies by using secretion vectors; examples include GM-CSF (Walter et al., 1992) and the extracellular domain of the human growth hormone receptor (deVos et al., 1992). The aforementioned proteins have been crystallized and their structures determined by X-ray crystallography, supporting the view that the structural integrity and conformation of the proteins were not affected by the lack of glycosylation and the respective preparative histories.

If a glycoprotein of interest is available from a eukaryotic recombinant expression system or if the natural protein is available, then before investing time with *E. coli* expression it may be worthwhile to determine whether the protein can be denatured and refolded in vitro. Pilot experiments can be carried out on intact protein and on protein enzymatically deglycosylated with glycosidases and, if disulfides are present, with and without reduction. Of course, if the protein can be secreted to the periplasm, aggregation and the necessity for in vitro folding may be avoided.

The production of deglycosylated proteins in *E. coli* expression systems for in vitro biochemical and structural studies is obviously of great value; however, the proteins may not always be suitable for in vivo studies due to low biological activity. Compared to authentic proteins, nonglycosylated variants can have a reduced circulatory lifetime and can exhibit increased immunogenicity and protease sensitivity (Rasmussen, 1992). Hodgson (1993) reviews the various expression systems used by biotechnology companies and lists the therapeutic proteins in development, many of which are produced in *E. coli* and which are normally glycosylated.

PROTEIN HANDLING

Storing Purified Proteins

Purified protein should be filter sterilized prior to storage. Millex-GV 0.22-μm filters (Millipore) are based on hydrophilic membranes with low binding capacities and are recommended for most proteins. Proteins are best stored at –80°C or may be stored on ice; freezing at –20°C is not recommended. Rapid freezing in small aliquots using dry ice/ethanol mixtures is preferred to slow freezing in the refrigerator. The addition of sucrose or glycerol often increases protein stability during storage and during freezing and thawing cycles (Arakawa and Timasheff, 1985; Timasheff and Arakawa, 1989). Lyophilization is best for long-term storage; however, care should be taken in choosing the protein solvent (Franks, 1993).

Promoting Protein Solubility and Stability

If the recombinant protein contains reactive unpaired sulfhydryl groups in the native conformation, 1 to 5 mM DTT should be included in the column buffers during purification. However, reductant should not be used gratuitously, as the native protein may contain intra- or intermolecular disulfide bonds, disruption of which can reduce the stability and solubility of the protein. Reductants should be included, for example, during gel filtration if dimers or higher aggregates need to be converted to active monomeric protein. The presence of intermolecular (and occasionally intramolecular) disulfide bonds can be determined analytically by SDS-PAGE under nonreducing conditions (*UNIT 6.5*): proteins are pretreated sequentially with iodoacetamide (to prevent artificial disulfide exchange) and SDS in the absence of reductant. The use of reductants can best be rationalized once the native protein has been characterized.

EDTA (1 to 5 mM) is often included in buffers to remove heavy metals that can catalyze oxidative processes and to inhibit certain proteases. It should be noted that EDTA will bind to anion exchangers (Scopes, 1994).

Other components often added to buffers to promote protein solubility during purification

ECU IASH CENTRAL UNIVERSITY LIBRARY

include nonionic or zwitterionic detergents, low concentrations of urea (1 to 2 M), and salt (0.5 to 1 M NaCl). These additives are compatible with ion-exchange chromatography, except for high salt concentrations, which are compatible with hydrophobic-interaction chromatography (*UNIT 8.4*), affinity chromatography (Chapter 9), and gel-filtration chromatography (*UNIT 8.3*). Solvent pH is one of the most important variables for maintaining protein solubility; in general, proteins are least soluble at or near their isolectric points.

Preventing Contamination

Precautions to prevent contamination of the protein of interest are as follows:

1. To avoid cross-contamination, especially from other recombinant proteins, reserve one set of chromatography resins for the purification of each protein. If this is not possible, or if expensive prepackaged matrices are used, be sure to clean resins thoroughly after each use; check the manufacturer's recommendations and be aware of the chemical stability of the resin, especially for extremes of pH.
2. Store resins with preservatives (e.g., sodium azide) and avoid storage in phosphate buffers, which provide a good medium for bacterial growth.
3. To generate reproducible protocols using ion-exchange methods, monitor the pH and conductivity of all buffers and column effuents (the latter ideally in-line).
4. Avoid protein cross-contamination in concentration equipment such as stirred ultrafiltration cells with ultrafiltration membranes.
5. Keep pH and conductivity probes scrupulously clean, especially when used with solutions containing proteases. Likewise, observe care when using cuvettes for UV measurements.
6. Avoid rigorous stirring of protein solutions to prevent shear denaturation, and handle soft agarose-based column matrices carefully to prevent bead fragmentation.

Removing Pyrogens

Recombinant proteins used for in vivo studies should be free of endotoxins (pyrogenic lipopolysaccharide from the bacterial outer membrane). Pyrogens can be detected using the sensitive *Limulus* amoebocyte lysate (LAL) assay with kits available from Sigma and other suppliers, and they can be removed using polymin B immobilized on agarose (Detoxi-Gel from Pierce). For further discussion of pyrogens and nucleic acid contamination, see Walsh and Headon (1994b).

SCALE OF OPERATIONS AND AIMS OF PURIFICATION

Determining Scale

The amount of protein required and the level of purity will vary dramatically from laboratory to laboratory and study to study. The following guidelines will help in planning a stategy.

If a Coomassie blue-stained band corresponding to the expressed protein is observed on one-dimensional SDS-PAGE analysis of a whole-cell extract, then the protein constitutes at least 0.5% to 1% of the total protein. Wet *E. coli* cell paste contains ~10% to 15% protein by weight (reviewed by Neidhardt, 1987). If the level of expression is low to average (e.g., 5%), then 1 g wet weight cells will contain ~5 mg recombinant protein. Hence, a cell paste of 20 to 50 g (a typical yield from a 1- to 2-liter benchtop fermentation) will contain 100 to 250 mg recombinant protein, and often 2- to 5-fold more. (Shaker flask fermentations of equivalent volumes might yield 5% or 10% of these amounts.) Thus, for soluble proteins, or insoluble ones that can be refolded (with ≥5% yield), significant amounts of protein can be obtained from relatively small fermentations. For proteins secreted into the periplasm or medium, fermentations at larger scales may be required, as expression levels are usually considerably lower than for direct expression.

Deciding the Aims of the Purification

There are obviously many reasons, both scientific and commercial, for producing purified recombinant proteins. The development of laboratory-scale purification schemes that produce a single band of protein on SDS-polyacrylamide gels should not pose major problems given the relatively high abundance of recombinant proteins in cell extracts, and assuming that protein folding has been solved. Protein present at 1% of the total cell extract requires only a 100-fold purification compared to the several thousandfold sometimes required for the purification of nonrecombinant proteins (reviewed by Stein, 1991). However, far more time and expertise may be needed to develop protocols that produce purified recombinant proteins having the physical and chemical homogeneity required for clinical-grade materials and for structural determinations.

A single band on an SDS-polyacrylamide gel may exhibit many bands on an isoelectric

BURNS / METRAL UNIVERSITY LIBRARY

focusing gel (*UNIT 7.1*). Only after detailed characterization of the isolated protein will the nature of the chemical and physical microheterogeneities be understood well enough to either prevent their occurrence or rationalize how to fractionate unmodified from modified species.

For many investigators, a primary goal is to correlate the structure of a protein with its function (and vice versa). Many proteins produced by recombinant DNA technology are present only in trace amounts in nature (e.g., interferons and other cytokines; Ealick et al., 1991), and authentic material is not available for detailed molecular characterization. Knowledge of the three-dimensional structure allows a rational approach to protein engineering and the design of drugs that modulate the biological activity of the protein. The substitution, deletion, and insertion of residues allow a structure-function hypothesis to be tested and new, sometimes improved protein variants (or mutants) to be produced.

Structural determination by X-ray diffraction requires good-quality crystals. Working out crystallization conditions may take from a few milligrams (lucky) to a hundred milligrams or more (realistic). The protein itself must usually be physically and chemically homogeneous; the presence of small amounts of unrelated *E. coli* protein impurities may interfere with crystallization. In general, physical homogeneity is more critical than chemical homogeneity. Some of the methods used to establish physical and chemical homogeneity are discussed elsewhere (Chapter 7; Jones et al., 1994). Structure determination by the X-ray method can be greatly aided by in vivo labeling of recombinant proteins with heavy atom derivatives as described in *UNIT 5.3*.

Structural determination in solution by multidimensional NMR is presently limited to proteins <30 kDa (reviewed by Clore and Gronenborn, 1994). The sample demands can be as high as several hundred milligrams, and larger proteins (>10 kDa) must be uniformly labeled with various combinations of ^{2}H, ^{13}C, and ^{15}N (the usual requirements are for protein singly labeled with ^{15}N and for protein doubly labeled with ^{15}N and ^{13}C). Labeling is achieved by growing the bacteria in minimal medium containing one or more of the following stable (nonradioactive) isotopes: $^{15}NH_4Cl$ (sole nitrogen source), [^{13}C]glucose (sole carbon source), and $^{2}H_2O$ (*UNIT 5.3*).

For NMR purposes the recombinant protein must be homogeneous and soluble at 1 to 3 mM concentrations, preferably with solvents below pH 7.0 and at temperatures above 30°C. As measurements take many hours to complete, the presence of trace amounts of proteases can ruin the entire experiment. In addition, particular attention must be paid to maintaining solvent-accessible and reactive cysteines (unpaired) in the reduced state (usually by including DTT).

Low-resolution structural studies using various biophysical methodologies (Jones et al., 1994) can be made with less material (<1 to 10 mg). Proteins for spectroscopic studies should be at least 95% pure and previously fractionated on a gel-filtration column to remove aggregated and possibility misfolded variants.

Finally, with in vivo protein labeling scenarios as described for isotopes (*UNIT 5.3*), tryptophan in recombinant proteins can be replaced by 5-hydroxytryptophan by using an *E. coli* Trp auxotroph. Protein thus labeled has a strong absorbance at 310 nm that can be exploited in structure-function studies (Laue et al., 1993).

SPECIALIZED EQUIPMENT

Breaking and Fractionating Cells

For small- to medium-scale work on a regular basis, an Aminco French press (SLM Instruments) with a continuous-fill cell is recommended (*UNIT 6.2*). This equipment is also useful for breaking yeast cells. For large-scale work (>500 ml), the Manton-Gaulin-APV homogenizer (APV Gaulin) is recommended. For further processing of cells and cell lysates (e.g., *UNITS 6.2 & 6.3*), an ultrasonic homogenizer is required. An instrument with a 400-W (or higher) capacity is recommended (Branson or other vendor).

After low-speed centrifugation using standard preparative centrifuges, high-speed centrifugation is a convenient and rapid cleanup step before column chromatography (Fig. 6.1.3). With Beckman ultracentrifuges, the 45 Ti rotor is recommended. This six-place rotor has a maximum speed of 235,000 × *g*; with thick-walled polycarbonate tubes, its capacity is ~400 ml.

Chromatographing Proteins

Most chromatography is carried out at 4°C either in a cold room or, more conveniently, in a cold cabinet in the laboratory. The basic components of a chromatography system are as follows: column, column matrix, pumps, a gradient-making device, UV/visible or other detection system, and a fraction collector.

ECU IAS/CSURA - UNIVERSITY LIBRARY

These components can be bought as units such as Pharmacia Biotech's FPLC (small- to medium-scale work) and BioPilot systems (medium- to large-scale work), or they can be assembled using individual components from various vendors. Column matrices can be purchased prepacked or as bulk media that are packed in columns by the user. Ion-exchange separations, using standard low- to medium-pressure resins (agarose/dextran/cellulose–based), require at least one narrow (2.5-cm) and one wide (5.0-cm) column with adjustable flow adapters so that the resin height can be varied between 5 and 30 cm. Gel filtration requires columns with diameters of 1.25 and 2.5 cm (5 cm for larger scale work) and lengths of 60 to 100 cm. Simple gradient makers with capacities of 150 ml to 2 liters can be purchased from Life Technologies (formerly GIBCO/BRL).

Concentrating Proteins

Stirred ultrafiltration cells are recommended for laboratory-scale work. The cells range in size from 3 ml to 2 liters and are used in conjunction with variable molecular weight cutoff membranes (Amicon). For larger volumes, both Millipore and Amicon sell various systems (see respective manufacturer's literature). For smaller volumes (0.5 to 15 ml), centrifugational concentrators are available (Amicon and others). For a review of the equipment used for protein concentration, see Harris (1989).

Making Analytical Measurements

A protein purification laboratory should have a good scanning UV/visible spectrophotometer, ideally an instrument with computerized data collection and analysis. Hewlett Packard instruments with diode array detectors are recommended for most routine work. For laboratories specializing in purifying recombinant proteins from *E. coli*, access to a spectropolarimeter (e.g., Jasco J-720) will be helpful for monitoring and developing folding protocols.

LITERATURE CITED

Allet, B., Payton, M., Mattaliano, R.J., Gronenborn, A.M., Clore, G.M., and Wingfield, P.T. 1988. Purification and characterization of the DNA-binding protein Ner of bacteriophage Mu. *Gene* 65:259-268.

Amrein, K.E., Takacs, B., Steiger, M., Molnos, J., Flint, N.A., and Burn, P. 1995. Purification and characterization of recombinant human $p50^{csk}$ protein-tyrosine kinase form an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc. Natl. Acad. Sci. U.S.A.* 92:1048-1052.

Arakawa, T. and Timasheff, S.N. 1985. Theory of protein solubility. *Methods Enzymol.* 114:49-77.

Asenjo, J.A. and Patrick, I. 1990. Large-scale protein purification. *In* Protein Purification Applications: A Practical Approach (E.L.V. Harris and S. Angal, eds.) pp. 1-27. IRL Press, Oxford.

Atkinson, D.E., Clark, S.G., and Rees, D.C. 1989. Equilibrium and acid-base relationships. *In* Dynamic Models in Biochemistry: A Workbook of Computer Simulations Using Electronic Spreadsheets (D.S. Barkley, ed.), Chapter 3. N. Simonson & Company, Marina del Rey, Calif.

Beacham, I.R. 1979. Periplasmic enzymes in gram-negative bacteria. *Int. Biochem.* 10:877-883.

Ben-Bassat, A., Bauer, K., Chang, S.-Y., Myambo, K., Boosman, A., and Chang, S. 1987. Processing of the initiation methionine from proteins: Properties of the *E. coli* methionine aminopeptidase and its gene structure. *J. Bacteriol.* 169:751-757.

Bibi, E. and Béjà, O. 1994. Membrane topology of multidrug resistant protein expressed in *E. coli*. *J. Biol. Chem.* 31:19910-19915.

Bowden, G.A., Paredes, A.M., and Georgiou, G. 1991. Structure and morphology of protein inclusion bodies in *E. coli*. *Bio/Technology* 9:725-730.

Buchner, J. and Rudolph, R. 1992. Renaturation and characterization of recombinant Fab fragments produced in *Escherichia coli*. *Bio/Technology* 9:157-162.

Burgess, R.R. and Jendrisak, J.J. 1975. A procedure for the rapid, large-scale purification of *E. coli* DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *J. Biol. Chem.* 14:4634-4638.

Cheng, Y.-S.E., Kwoh, D.Y., Kwoh, J.T., Soltvedt, B.C., and Zipser, D. 1981. Stabilization of degradable protein by its overexpression in *E. coli*. *Gene* 14:121-130.

Cheng, Y.-S.E., McGowan, M.H., Kettner, C.A., Schloss, J.V., Erickson-Viitanen, S., and Yin, F.H. 1990. High synthesis of recombinant HIV-1 protease and the recovery of active enzyme from inclusion bodies. *Gene* 87:243-248.

Cleary, S., Mulkerrin, M.G., and Kelley, R.F. 1989. Purification and characterization of tissue plasminogen activator kringle-2 domain expressed in *E. coli*. *J. Biol. Chem.* 28:1884-1891.

Cleland, J.L., Builder, S.E., Swartz, J.R., Winkler, M., Chang, J.Y., and Wang, D.I.C. 1992. Polyethylene glycol enhanced protein refolding. *Bio/Technology* 10:1013-1019.

Clore, G.M. and Gronenborn, A.M. 1994. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* 239:249-363.

Creighton, T.E. 1984. Disulfide bond formation in proteins. *Methods Enzymol.* 107:305-329.

RUTGERS STATE UNIVERSITY CENTRAL LIBRARY

Creighton, T.E. 1993. Proteins: Structures and Molecular Properties, 2nd ed., pp. 292-296. Freeman, New York.

Dale, G.E., Broger, C., Langen, H., D'Arcy, A., and Struber, D. 1994. Improving protein stability through rationally designed amino acid replacements: Solubilization of the trimethoprim-resistant type S1 dihydrofolate reductase. *Prot. Eng.* 7:933-939.

Danley, D.E., Strick, C.A., James, L.C., Lanzetti, A.J., Otterness, I.G., Grenett, H.E., and Fuller, G.M. 1991. Identification and characterization of a C-terminally extended form of recombinant murine IL-6. *FEBS Lett.* 283:135-139.

Darby, N.J. and Creighton, T.E. 1990. Folding proteins. *Nature* 344:715-716.

deVos, A.M., Ultsch, M., and Kossiakoff, A.A. 1992. Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science* 255:306-312.

Diederichs, K., Boone, T., and Karplus, A. 1991. Novel fold and putative receptor binding site of granulocyte-macrophage colony stimulating factor. *Science* 254:1779-1782.

DiRienzo, J.M., Nakamura, K., and Inouye, M. 1978. The outer membrane proteins of gram-negative bacteria: Biosynthesis, assembly and function. *Annu. Rev. Biochem.* 47:481-532.

Ealick, S.E., Cook, W.J., Senadhi, V.-K., Carson, M., Nagabhushan, T.L., Trotta, P.P., and Bugg, C.E. 1991. Three-dimensional structure of recombinant human interferon-γ. *Science* 252:698-702.

Ellis, R.J. 1994. Role of chaperones in protein folding. *Curr. Opin. Struct. Biol.* 4:117-122.

Franks, F. 1993. Storage stabilization of proteins. *In* Protein Biotechnology (F. Franks, ed.) pp. 486-531. Humana Press, Clifton, N.J.

Goff, S.A. and Goldberg, M.E. 1985. Production of abnormal proteins in *E. coli* stimulates transcription of *lon* and other heat shock genes. *Cell* 41:587-595.

Goldberg, M.E. 1991. Investigating protein conformation dynamics and folding with monoclonal antibodies. *Trends Biochem. Sci.* 16:358-362.

Goldberg, M.E., Rudolph, R., and Jaenicke, R. 1991. A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. *Biochemistry* 30:2790-2797.

Grunfeld, H., Patel, A., Shatzman, A., and Nishikawa, A.H. 1992. Effector-assisted refolding of recombinant tissue-plasminogen activator produced in *Escherichia coli*. *Appl. Biochem. Biotechnol.* 33:117-138.

Harris, E.L. 1989. Concentration of the extract. *In* Protein Purification Methods: A Practical Approach (E.L.V. Harris and S. Angal, eds.) pp. 125-172. IRL Press, Oxford.

Hayano, T., Takahashi, N., Kato, S., Maki, N., and Suzuki, M. 1991. Two distinct forms of peptidylprolyl-*cis-trans*-isomerase are expressed separately in periplasmic and cytoplasmic compartments of *Escherichia coli* cells. *Biochemistry* 30:3041-3048.

Helenius, A. 1994. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol. Biol. Cell.* 5:253-265.

Heppel, L.A. 1967. Selective release of enzymes from bacteria. *Science.* 156:1451-1455.

Hodgson, J. 1993. Expression systems: A user's guide. *Bio/Technology* 11:887-893.

Holland, I.B., Kenny, B., Steipe, B., and Pluckthun, A. 1990. Secretion of proteins in *E. coli*. *Methods Enzymol.* 182:132-143.

Hopkins, T.R. 1991. Physical and chemical cell disruption for the recovery of intracellular proteins. *In* Purification and Analysis of Recombinant Proteins (R. Seetharam and S.K. Sharma, eds.) pp. 57-83. Marcel Dekker, New York.

Hwang, C., Sinskey, A.J., and Lodish, H.F. 1992. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* 257:1496-1502.

Jaenicke, R. 1993. Role of accessory proteins in protein folding. *Curr. Opin. Struct. Biol.* 3:104-112.

Jaenicke, R. and Rudolph, R. 1989. Folding proteins. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 191-223. IRL Press, Oxford.

Janson, J.-C. and Rydén, L. 1989. Protein Purification: Principles, High Resolution Methods, and Applications. VCH Publishers, New York.

Johnson, B.H. and Hecht, M. H. 1994. Recombinant proteins can be isolated from *E. coli* by repeated cycles of freezing and thawing. *Bio/Technology* 12:1357-1360.

Jones, C., Mulloy, B., and Thomas, A.H. (eds). 1994. Microscopy, optical spectroscopy, and macroscopic techniques. *Methods Mol. Biol.* 22:1-245.

Kaback, H.R. 1971. Bacterial membranes. *Methods Enzymol.* 22:99-120.

Kane, J.K. and Hartley, D.L. 1991. Properties of recombinant protein-containing inclusion bodies in *E. coli*. *In* Purification and Analysis of Recombinant Proteins (R. Seetharam and Sharma, eds.) pp. 121-145. Marcel Dekker, New York.

Kelley, W.S. and Stump, K.H. 1979. A rapid procedure for isolation of large quantities of *E. coli* DNA polymerase 1 utilizing a λ *polA* transducing phage. *J. Biol. Chem.* 254:3206-3210.

Kern, G., Kern, D., Jaenicke, R., and Seckler, R.L. 1993. Kinetics of folding and association of differentially glycosylated variants of invertase from *Saccharomyces cerevisiae*. *Protein Sci.* 2:1862-1868.

Kiefhaber, T., Rudolph, R., Kohler, H.-H. and Buchner, J. 1991. Protein aggregation in vitro and in vivo: A quantitative model of the kinetic competition between folding and aggregation. *Bio/Technology* 9:825-829.

BCU IASH CENTRAL UNIVERSITY LIBRARY

Kohno, T., Carmichael, D.F., Sommer, A., and Thompson, R.C. 1990. Refolding of recombinant proteins. *Methods Enzymol.* 185:187-195.

Laue, T.M., Shah, B.D., Ridgeway, T.M., and Pelletier, S.L. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science (S.E. Harding, A.J. Rowe, and J.C. Horton, eds.) pp. 90-125. Royal Society for Chemistry and Cambridge University Press, Cambridge.

Laue, T.M., Senear, D.F., Eaton, S., and Ross, J.B.A. 1993. 5-Hydroxytryptophan as a new intrinsic probe for investigating protein-DNA interactions by analytical ultracentrifugation. Study of the effect of DNA on the self-assembly of the bacteriophage λ *cI* repressor. *J. Biol. Chem.* 32:2469-2472.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology* 11:187-193.

London, J., Skrzynia, C., and Goldberg, M.E. 1974. Renaturation of *Escherichia coli* tryptophanase after exposure to 8 M urea. *Eur. J. Biochem.* 47:409-415.

Lu, H.S., Fausset, P.R., Sotos, L.S., Clogston, C.L., Rohde, M.F., Stoney, K.S., and Herman, A.C. 1993. Isolation and characterization of three recombinant human granulocyte colony stimulating factor His to Gln isoforms produced in *E. coli*. *Protein Expression Purification* 4:465-472.

Marston, F.A.O. and Hartley, D.L. 1990. Solubilization of protein aggregates. *Methods Enzymol.* 182:264-276.

Matthew, J.B., Friend, S.H., Botelho, L.D., Lehman, L.D., Hanania, G.I., and Gurd, F.R.H. 1978. Discrete charge calculations of potentiometric titrations for globular proteins. *Biochem. Biophys. Res. Commun.* 81:416-421.

Maurizi, M.R. 1992. Proteases and protein degradation in *Escherichia coli*. *Experientia* 48:178-201.

Milburn, M.V., Hassel, A.M., Lambert, M.H., Jordon, S.R., Proudfoot, A.E.I., Graber, P., and Wells, T.N.C. 1993. A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature* 363:172-176.

Mildner, A.M., Rothrock, D.J., Leone, J.W., Bannow, C.A., Lull, J.M., Reardon, I.M., Sarcich, J.L., Howe, W.J., Tomich, C.-S.C., Smith, C.W., Heinrikson, R.L., and Tomasselli, A.G. 1994. The HIV-1 protease as enzyme and substrate: Mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. *Biochemistry* 33:9405-9413.

Miller, C.G., Strauch, K.L., Kurral, A.M., Miller, J.L., Wingfield, P.T., Mazzei, G.J., Werlen, R.C., Graber, P., and Movva, N.R. 1987. N-Terminal methionine-specific peptidase in *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. U.S.A.* 84:2718-2772.

Miller, H.I., Henzel, W.J., Ridgway, J.B., Kuang, W.-J., Chrisholm, V., and Liu, C.-C. 1989. Cloning and expression of a yeast ubiquitin-protein cleaving activity in *E. coli*. *Bio/Technology* 7:698-704.

Missiakas, D., Georgopoulos, C., and Raina, S. 1994. The *Escherichia coli dsbC xprA* gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J.* 13:2013-2020.

Nagata, K., Kikuchi, N., Ohara, O., Teraoka, H., Yoshida, N., and Kawade, Y. 1986. Purification and characterization of recombinant murine immune interferon. *FEBS Lett.* 205:200-204.

Neidhardt, F.C. 1987. Chemical composition of *Escherichia coli*. *In Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology (F.C. Neidhardt, ed.) pp. 3-6. American Society for Microbiology, Washington, D.C.

Orsini, G. and Goldberg, M.E. 1978. The renaturation of reduced chymotrypsin A in guanidine HCl. *J. Biol. Chem.* 253:3453-3458.

Patel, D. 1993. Chromatographic fractionation media. *In* Biochemistry Labfax (J.A.A. Chambers and D. Rickwood, eds.) pp. 49-68. BIOS Scientific Publishers and Academic Press, Oxford.

Puri, N.K., Crivelli E., Cardamome, M., Fiddes, R., Bertoloini, J., Ninham, B., and Brandon, M.R. 1992. Solubilization of growth hormone and other recombinant proteins from *Escherichia coli* by using a cationic surfactant. *Biochem. J.* 285:871-879.

Rasmussen, J.R. 1992. Effect of glycosylation on protein function. *Curr. Opin. Struct. Biol.* 2:682-686.

Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *Bio/Technology.* 7:1141-1147.

Schmid, F.X. 1989. Spectral methods of characterizing protein conformation and conformational change. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 251-285. IRL Press, Oxford.

Schneider, C., Newman, R.A., Sutherland, D.R., Asser, U., and Greaves, M.F. 1982. A one step purification of membrane proteins using a high affinity immunomatrix. *J. Biol. Chem.* 257:10766-10769.

Scopes, R.K. 1994. Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.

Sherman, P.A. and Fyfe, J.A. 1990. Human immunodeficiency virus integration protein expressed in *E. coli* possesses selective DNA cleaving activity. *Proc. Natl. Acad. Sci. U.S.A.* 87:5119-5123.

Sherman, F., Stewart, J.W., and Tsunasawa, S. 1985. Methionine or not methionine at the beginning of a protein. *Bioessays* 3:27-31.

Shire, S.J., Bock, L., Ogez, J., Builder, S., Kleid, D., and Moore, D.M. 1984. Purification and immunogenicity of fusion VP1 protein of foot and

PRESS / CONTRA UNIVERSITY LIBRARY

mouth disease virus. *Biochemistry* 23:6474-6480.

Sinha, D., Bakhshi, M., and Vora, R. 1994. Ligand binding assay with recombinant protein refolded on an affinity matrix. *BioTechniques* 17:509-514.

Skerra, A., Pfitzinger, I., and Pluckthun, A. 1991. The functional expression of antibody Fv fragments in *Escherichia coli*: Improved vectors and a generally applicable purification technique. *Bio/Technology* 9:273-278.

Stark, G.R., Stein, W.H., and Moore, S. 1960. Reactions of cyanate present in aqueous urea with amino acids and proteins. *J. Biol. Chem.* 235:3177-3181.

Stein, S. 1991. Isolation of natural proteins. *In* Fundamentals of Protein Biotechnology (S. Stein, ed.) pp. 137-160. Marcel Dekker, New York.

Sutcliffe, J.G., Shinnick, T.M., Green, N., and Lerner, R.A. 1983. Antibodies that react with predetermined sites on proteins. *Science* 219:660-665.

Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* 23:122-275.

Thatcher, D.R. 1990a. Large-scale production of hematopoietic growth factors. *Immunology Ser.* 49:329-357.

Thatcher, D.R. 1990b. Recovery of therapeutic proteins from inclusion bodies: Problems and process strategies. *Biochem. Soc. Trans.* 18:234-235.

Thatcher, D.R. and Panayotatos, N. 1986. Purification of recombinant IFN-α2. *Methods Enzymol.* 119:166-177.

Timasheff, S.N. and Arakawa, T. 1989. Stabilization of protein structure by solvents. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 331-345. IRL Press, Oxford.

Uhlén, M., Forberg, G., Moks, T., Hartmanis, M., and Nilsson, B. 1992. Fusion proteins in biotechnology. *Curr. Opin. Biotechnol.* 3:363-369.

Ultsch, M., deVos, A.M., and Kossiakoff, A.A. 1991. Crystals of the complex between human growth hormone and the extracellular domain of its receptor. *J. Mol. Biol.* 222:865-868.

Walsh, G. and Headon, D.R. 1994a. Downstream processing of protein products. *In* Protein Biotechnology, pp. 39-117. John Wiley & Sons, New York.

Walsh, G. and Headon, D.R. 1994b. Therapeutic proteins: Special considerations. *In* Protein Biotechnology, pp. 118-162. John Wiley & Sons, New York.

Walter, H. and Johansson, G. 1986. Partitioning in aqueous two-phase systems: An overview. *Anal. Biochem.* 155:215-242.

Walter, M.R., Cook, W.J., Ealick, S.E., Nagabhushan, T.L., Trotta, P.P., and Bugg, C.E. 1992. Three-dimensional structure of human recombinant granulocyte-macrophage colony stimulating factor. *J. Mol. Biol.* 224:1075-1085.

Watanabe, E., Tsoko, S., and Asenjo, J.A. 1994. Selection of chromatographic protein purification operations based on physicochemical properties. *Ann. N.Y. Acad. Sci.* 721:348-364.

Wetlaufer, D.B. 1962. Ultraviolet spectra of proteins and amino acids. *Adv. Prot. Chem.* 17:303-390.

Wetlaufer, D.B. 1984. Nonenzymatic formation and isomerization of protein disulfides. *Methods Enzymol.* 107:301-304.

Wetlaufer, D.B., Branca, P.A., and Chen, G.-X. 1987. The oxidative folding of proteins by disulfides plus thiol does not correlate with redox potential. *Protein Eng.* 1:141-146.

Wetzel, R. 1992. Principles of protein stability. Part 2—Enhanced folding and stabilization of proteins by suppression of aggregation in vitro and in vivo. *In* Protein Engineering: A Practical Approach (A.R. Rees, M.J.E. Sternberg, and R. Wetzel, eds.) pp. 191-216. IRL Press, Oxford.

Wingfield, P.T., Mattaliano, R.J., MacDonald, H.R., Craig, S., Clore, G.M., Gronenborn, A.M., and Schmeissner, U. 1987a. Recombinant-derived interleukin 1α stabilized against specific deamidation. *Protein Eng.* 1:413-417.

Wingfield, P.T., Graber, P., Rose, K., Simona, M.G., and Hughes, G.J. 1987b. Chromatofocusing of N-terminally processed forms of proteins: Isolation and characterization of two forms of interleukin 1β and bovine growth hormone. *J. Chromatogr.* 387:291-300.

Wingfield, P.T., Payton, M., Graber, P., Rose, K., Dayer, J.-M., Shaw, A.R., and Schmeissner, U. 1987c. Purification and characterization of human interleukin 1α produced in *Escherichia coli*. *Eur. J. Biochem.* 165:537-541.

Wingfield, P.T., Stahl, S.J., Payton, M.A., Venkatesan, S., Misra, M., and Steven, A. 1990. HIV-1 Rev expressed in recombinant *Escherichia coli*: Purification polymerization and conformational properties. *Biochemistry* 30:7527-7534.

Wingfield, P.T., Stahl, S.J., Williams, R.W., and Steven, A.C. 1995. Hepatitis core antigen produced in *E. coli*: Conformational analysis, and in vitro assembly. *Biochemistry* 34:4919-4932.

Yarranton, G.T. and Mountain, A. 1992. Expression of proteins in prokaryotic systems—Principles and case studies. *In* Protein Engineering: A Practical Approach (A.R. Rees, M.J.E. Sternberg, and R. Wetzel, eds.) pp. 303-324. IRL Press, Oxford.

Zapun, A., Missiakas, D., Raina, S., and Creighton, T.E. 1995. Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry* 34:5075-5089.

Zardeneta, G. and Horowitz, P.M. 1994. Detergent, liposome, and micelle-assisted protein refolding. *Anal. Biochem.* 223:1-6.

Contributed by Paul T. Wingfield
National Institutes of Health
Bethesda, Maryland

BCU IAS LIBRARY

Preparation of Soluble Proteins from *Escherichia coli*

Once a suitable protein expression system involving *Escherichia coli* is developed and optimized (*UNITS 5.1 & 5.2*), large-scale production of recombinant proteins (*UNIT 5.3*) generates large quantities of culture material from which the protein of interest must be purified. Harvesting (*UNIT 5.3*) produces cell concentrate or culture medium, depending on the subcellular localization of the protein. Cell paste is the starting material for purification of proteins expressed in soluble form inside cells, such as interleukin 1β (IL-1β). Human IL-1β is a 153-residue (17.4-kDa) protein cytokine of biomedical importance that plays a central role in immune and inflammatory responses. Purification of human IL-1β is used as an example of the preparation of soluble proteins from *E. coli*.

Bacteria containing IL-1β are lysed, and the resulting supernatant is clarified to remove ribosomes and other particulate matter. The sample is then applied to an anion-exchange column to separate recombinant IL-1β from cellular contaminants, such as *E. coli* proteins, nucleic acids, and lipopolysaccharides. The sample is further purified through salt precipitation and cation-exchange chromatography, then concentrated. Finally, the IL-1β protein is applied to a gel-filtration column to separate it from remaining higher- and lower-molecular-weight contaminants, the purified protein is stored frozen or is lyophilized.

The purification protocol described is typical for a protein that is expressed in fairly high abundance (i.e., >5% total protein) and accumulates in a soluble state. With these expression levels, only about a 20-fold overall purification is required to obtain pure protein. Therefore, conventional chromatographic methods can be used, and normally only three or four purification stages are required (see Table 6.2.1 for an outline of the procedure as applied to IL-1β, including time considerations). The process can be shortened somewhat through the use of the Pharmacia Biotech BioPilot or FPLC systems (see Time Considerations).

SDS-PAGE (*UNIT 10.1*) is used to monitor column fractions for the presence of IL-1β, which is detected as a stained band (*UNIT 10.5*) at the expected 17.4-kDa location. This is common practice in recombinant protein purification, as the specific assays for many proteins, including IL-1β, are relatively complex. The original expression and fermentation experiments involving IL-1β (Wingfield et al., 1986) followed procedures similar to those given in *UNIT 5.3*.

CAUTION: Avoid direct contact with IL-1β-containing solutions. Trace material from aerosols or from hand contact can cause severe inflammation of the eyes. Safety glasses and gloves should be worn.

PURIFICATION OF A PROTEIN EXPRESSED IN *ESCHERICHIA COLI* IN A SOLUBLE STATE: INTERLEUKIN 1β

BASIC PROTOCOL

Materials

DEAE Sepharose CL-4B resin (Pharmacia Biotech)
Anion-exchange buffer (see recipe)
0.26% (w/v) sodium hypochlorite/70% ethanol *or* 5% (v/v) bleach (e.g., Clorox)/70% ethanol
E. coli cells (~50 g wet weight) from fermentation (*UNIT 5.3*) containing IL-1β
Lysis buffer (see recipe)
Bovine pancreas DNase I and RNase A (Worthington; optional, for reducing solution viscosity)

Copyright © 1995 by John Wiley & Sons, Inc.

BOWLES CENTRAL LIBRARY UNIVERSITY

2 N sodium hydroxide
Ammonium sulfate, ground with mortar and pestle
Cation-exchange buffer (see recipe)
CM Sepharose CL-4B (Pharmacia Biotech)
Cation-exchange buffer/250 mM NaCl (see recipe)
Tris base
Gel-filtration buffer (see recipe)
Ultrogel AcA54 gel-permeation resin (BioSepra)
Lyophilization buffer (see recipe; optional)

2- or 3-liter sintered glass funnel with fritted disc (coarse porosity) and 5-liter filter flask
Chromatography columns (preferably glass) with adjustable flow adapters: one (or optionally two) 5 × 50 cm and one 2.5 × 100 cm (Pharmacia Biotech, Amicon, or equivalent)
RK50 packing reservoir (Pharmacia Biotech)
Peristaltic pump, UV monitor, and fraction collector (Pharmacia Biotech or equivalent)
16 × 150–mm culture tubes
40-ml French pressure cell and rapid-fill kit (SLM-AMINCO)
Aminco laboratory press (SLM-AMINCO)
1-liter Waring commercial blender

Table 6.2.1 Outline of Interleukin β Purification[a]

Day	Steps (low pressure)	Steps (BioPilot or FPLC)
1	Preparation of DEAE Sepharose column (steps 1-3)	As for low pressure (steps 1-17)
	Cell breakage (steps 4-8)	
	Clarification of lysate (steps 9 and 10)	
	DEAE Sepharose chromatography (step 11)	
	SDS-PAGE of DEAE Sepharose fractions (step 12)	
	$(NH_4)_2SO_4$ fractionation (steps 13-15)	
	Dialysis (steps 16 and 17)	
2	CM Sepharose chromatography (steps 18-22)	Fast Flow (substitute for CM Sepharose; steps 18-21)
	SDS-PAGE of CM Sepharose fractions (step 22)	SDS-PAGE of Fast Flow fractions (steps 22)
		Concentration (step 23a)
		Superdex 75 gel filtration (substitute for Ultrogel; steps 24-26)
3	Concentration (step 23a or 23b)	SDS-PAGE of Superdex 75 fractions (step 26)
	Ultrogel gel filtration (steps 24-26)	Concentration (step 27)
4	SDS-PAGE of gel-filtration fractions (step 26)	
	Concentration (step 27)	

[a]Step numbers in parentheses refer to the Basic Protocol. Chromatography materials as well as the BioPilot and FPLC systems are from Pharmacia Biotech.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

250, 500, and 1000-ml stainless steel beakers
Ice bucket, ~4 liter
Tissue-grinder homogenizer (Polytron Model PT 10/35, Brinkmann)
Ultrasonic homogenizer, ≥400 W, with sound enclosure (Branson or equivalent)
Preparative centrifuge: Beckman J2-21M
Rotors for preparative centrifuge: Beckman JA-14 (capacity 6 × 250 ml) or JA-20 (capacity 8 × 50 ml)
Ultracentrifuge: Beckman Optima XL-90
Rotors for ultracentrifuge: Beckman 45Ti (capacity 6 × 100 ml) or 35Ti (capacity 6 × 94 ml)
Conductivity meter (Radiometer America)
Spectra/Por 1 dialysis tubing (Spectrum)
Gradient maker: Model 2000 (working volume 0.65 to 2 liters; Life Technologies)
200- or 400-ml stirred ultrafiltration cell and Diaflo ultrafilter PM10 or YM3 membranes (Amicon; optional)
Millex-GV 0.22-µm-pore-size filter units (Millipore)
10- or 20-ml syringe

Additional materials and equipment for SDS-PAGE (*UNIT 10.1*) and dialysis (*APPENDIX 3B*)

NOTE: All protocol steps are carried at 4°C unless otherwise stated. Forces for centrifugation steps refer to the maximum × *g* (i.e., centrifugal force at the bottom of the tubes).

Prepare anion-exchange column

1. Pour 400 to 500 ml DEAE Sepharose CL-4B ion-exchange resin into a sintered-glass funnel and wash with several liters water followed by 1 liter anion-exchange buffer (pH 8.5). Measure the conductivity of the starting buffer and eluted buffer to make sure they are the same before proceeding to the next step.

   The resin is supplied in 500-ml bottles as a slurry in 20% ethanol. When washing the resin, do not allow it to run dry on the filter funnel. Laboratory vacuum (e.g., water aspirator) is adequate for filtering.

2. Suspend the washed resin in anion-exchange buffer to 75% settled gel/25% buffer by volume, per manufacturer's recommendations. Degas in a filter flask and pour into a 5 × 50–cm chromatography column fitted with a filling reservoir.

   After settling, the height of the resin should be ~20 to 25 cm (390 to 490 ml packed resin). For details on packing columns, see UNIT 8.4. Because the solubility of gases decreases with increases in temperature, it is usual practice to pack the column at room temperature and then run it in a cold room or cold box.

3. Elute column with anion-exchange buffer at 100 to 150 ml/hr using a peristaltic pump. Make sure there is no compression of column contents. Monitor the absorbance of the effluent at 260 or 280 nm with a UV detector. Collect ~15-ml fractions in 16 × 50–mm culture tubes using a fraction collector. Check that the pH and conductivity of the column effluent are the same as the for anion-exchange buffer applied to the column (this indicates that the column matrix is correctly equilibrated).

   The bed height should not change significantly once the column is packed. Compression indicates that the pressure applied to the column is too high (see manufacturer's recommendations for maximum flow rates).

Break cells with a French press

4. Clean bench areas that may come in direct contact with cells with 0.26% sodium hypochlorite/70% ethanol.

BU ARTICLE CENTRAL UNIVERSITY LIBRARY

5. Assemble the French pressure cell and and chill to ~4°C either by incubation in ice or by refrigeration. Install the cell (first dried with paper towels if necessary) in the Aminco laboratory press.

It is important to cool the equipment because pressurizing will generate heat. The 20K rapid-fill French pressure cell (1-in.-diameter piston) has a capacity of 40 ml and can be continuously filled while installed on the press. Before using the pressure cell, replace the nylon ball at the end of the flow valve assembly or, at the very least, check it for distortion.

For small-scale work, a miniature French pressure cell (3/8-in.-diameter piston) with a 3.7-ml capacity is available.

6. Suspend thawed *E. coli* cells (~50 g wet weight) with 150 ml lysis buffer using a Waring blender. Place the suspension in a stainless steel beaker and homogenize with the Polytron tissue-grinder homogenizer until clumps are no longer detected.

IMPORTANT NOTE: *Wear disposable gloves and safety glasses while working with E. coli. The high-pressure homogenization may generate aerosols.*

The E. coli cells are stored frozen at –80°C as a flattened paste in heat-sealable plastic bags (UNIT 5.3). The cells are thawed at room temperature. Complete suspension of the cells with the blender is important, as any visible clumps of bacteria will block the French pressure cell. A clogged cell may have to be disassembled to clear the blockage.

7. Lyse the cells with two passes through the French press operated at 16,000 to 18,000 lb/in^2 (with the high-ratio setting, pressure gauge readings between 1011 and 1135). Chill the cell suspension to 4°C after each pass through the pressure cell by incubation on ice.

When filling the pressure cell, avoid drawing air into the cylinder to prevent foaming.

If a French press is not available, the cells can be broken by including 200 µg/ml lysozyme (Worthington) and 0.05% (w/v) sodium deoxycholate (Calbiochem) in the lysis buffer and incubating cells ~20 min at 20° to 25°C with intermittent homogenization using the tissue grinder (Burgess and Jendrisak, 1975). Cell breakage by lysozyme treatment and sonication is described in Basic Protocol 3 of UNIT 6.5.

8. Place the suspension (contained in a steel beaker) on an ice bath and, using an ultrasonic homogenizer, sonicate 5 min at full power with 50% duty cycle (on for 0.5 sec then off for 0.5 sec). While sonicating, stir the suspension using a magnetic stirrer.

IMPORTANT NOTE: *Wear sound-protection earmuffs to protect ears from ultrasonic noise. Because sonication will generate some aerosol, use the sonicator in a microbiological hood if possible.*

High viscosity reduces the rate of sedimentation of the various contaminating cellular material and thus longer (sometimes much longer) centrifugation times are required. Sonication reduces the viscosity of the suspension prior to centrifugation by shearing the released DNA and RNA. The viscosity can also be reduced by digesting the lysate 15 to 30 min at 4° to 10°C with bovine pancreas DNase I (25 to 50 µg/ml) and RNase A (50 µg/ml). If the nucleases are used, EDTA in the lysis buffer should be replaced by 5 mM $MgCl_2$ as DNase requires Mg^{2+}.

Clarify the lysate

9. Transfer sample from the beaker to centrifuge bottles and centrifuge the cell lysate 40 min at 22,000 × *g* (e.g., in a Beckman J2-21M preparative centrifuge at 12,000 rpm using JA-14 rotor or at 13,500 rpm using JA-20 rotor), 4°C. Decant the supernatants, pool, and recentrifuge 90 min at ~100,000 × *g* (30,000 rpm in Beckman Optima XL-90 ultracentrifuge using Ti45 rotor), 4°C.

Low-speed centrifugation removes unbroken cells and large cellular debris. High-speed centrifugation removes smaller particles such as ribosomes and membrane

PCU IASI / CENTRAL UNIVERSITY LIBRARY

vesicles; the Beckman 70Ti rotor (capacity 8 × 39 ml) can be used in the ultracentrifuge for smaller-scale work. Clarification of the lysate can also be carried out by salt fractionation (see Background Information, section on determining solubility, for further details).

Pellets are usually discarded immediately; but see Critical Parameters and Troubleshooting, section on protein purification for further comments.

10. Dilute the supernatant from step 6 (~160 ml) 1:2 (three-fold) with anion-exchange buffer and adjust to pH 8.5 (if necessary) with 2 N NaOH. Using a conductivity meter, measure the conductivity of the diluted supernatant. If it is higher than 5.0 to 5.3 mS/cm, reduce by dilution with water.

The conductivity of the protein solution is carefully adjusted to ensure that the proteins (in this case contaminants) are bound to the matrix. Too high an ionic strength will reduce or prevent binding.

The DEAE Sepharose column is normally prepared (steps 1 to 3) before lysis of the cells is initiated. If there are any delays in applying sample to column, store the clarified lysate at 0° to 4°C (for example, in a covered beaker or flask embedded in an ice bucket). The same applies to the other chromatographic stages.

Chromatograph cleared lysate on anion-exchange resin

11. Apply the clarified lysate (480 to 500 ml) to the DEAE Sepharose column (step 3) at 150 ml/hr and elute the column with anion-exchange buffer. Continue the elution, collecting 15-ml fractions, until the effluent absorbance is close to the baseline value.

In solutions above ~pH 7.0, IL-1β (pI 6.8) is negatively charged. Therefore, the protein should, in principle, bind to an anion exchanger (positively charged matrix) buffered at pH 8.5. In fact, under the conditions described, IL-1β is only weakly bound to the matrix, which allows for partial resolution from proteins that do not bind to the matrix. Much of the unbound protein is of high molecular weight (or highly aggregated) and is partially separated from IL-1β by the gel filtration effect of the matrix. IL-1β is therefore usually located in the latter two-thirds (~500 ml) of the column flowthrough volume.

12. Assay every second or third column fraction by SDS-PAGE (*UNIT 10.1*).

See Figure 6.2.1A for an example of results from SDS-PAGE.

For rapid analysis, use precast gels or the Hoefer Pharmacia Phast system. Alternatively, if speed of purification is important, the entire flowthrough can be used, eliminating the need to analyze separate fractions. Most protein contaminants bound to the column can be removed by step elution with 1 M NaCl in column buffer. After use, the resin should be unpacked from the column and washed on a sintered-glass funnel with 1 liter of 2 M NaCl/0.5% (w/v) Triton X-100, followed by 10 liters water. If the resin is to be stored, suspend it in 5% ethanol or 5 mM sodium azide and store at 4°C. In order to avoid potential cross-contamination, dedicate the used resin for repeat purifications of IL-1β only.

Fractionate sample with ammonium sulfate

13. Pool IL-1β-containing fractions from the DEAE Sepharose column, record the volume (~500 ml), and transfer the solution to a 1-liter beaker (preferably stainless steel). Add 30.2 g $(NH_4)_2SO_4$ per 100 ml solution at 0°C (51.3% saturation or 2 M final concentration): add powdered $(NH_4)_2SO_4$ slowly over 30 min, mixing gently with a magnetic stirrer, then allow a further 30 min of mixing.

Methods for calculating the percent saturation of ammonium sulfate solutions, which specifically refer to 0°C, have been described by Wood (1976). When using solid ammonium sulfate, calculations must include volume increases on addition of the solid to fixed volumes (as carried out above).

BOULDER / GENERAL / UNIVERSITY LIBRARY

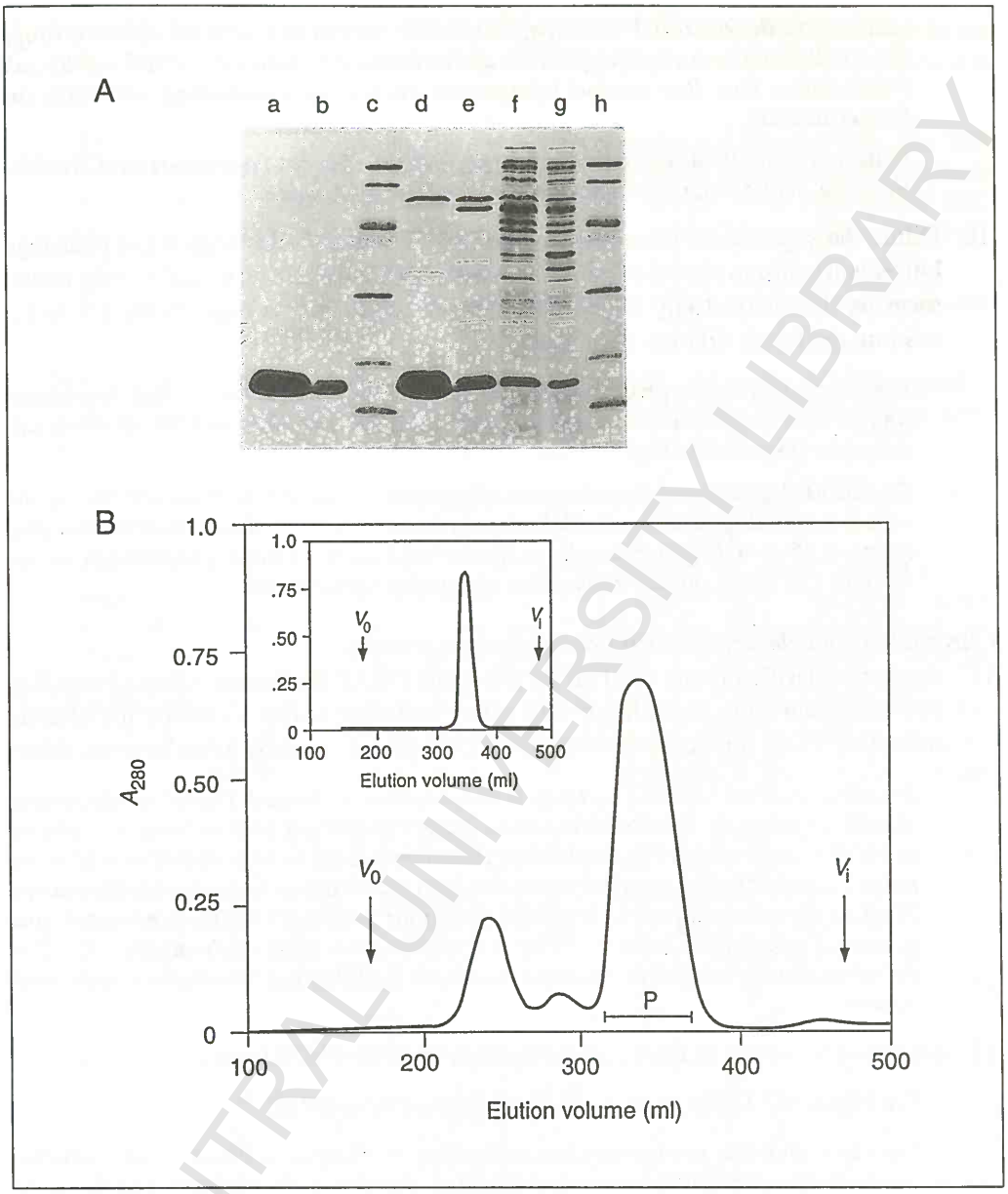


Figure 6.2.1 Purification of IL-1β. (**A**) SDS-PAGE analysis of samples at various stages. Analysis was conducted on a gel of dimensions 12 cm × 16 cm × 1.5 mm. Lane a, purified protein (100 µg loaded); lane b, purified protein (10 µg loaded); lane d, CM Sepharose pool (80%); lane e, DEAE Sepharose pool after ammonium sulfate fractionation (56%); lane f, high-speed supernatant (starting material for DEAE Sepharose column; 13.5%); lane g, cell lysate (12.0%). The percentages refer to specific IL-1β contents of the fractions determined by densitometric scanning of the Coomassie blue–stained gel lanes. Lanes c and h contain the following protein standards (low-range standards supplied by Bio-Rad) in order of increasing migration distance: phosphorylase *b* (97.4 kDa), bovine serum albumin (66.2 kDa), hen egg white ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), and hen white lysozyme (14.4 kDa). (**B**) Analysis of results from gel filtration on Ultrogel AcA54. The excluded volume (V_0) and the fully included volume (V_i) are indicated. Inset, analytical rechromatography of the protein from the pooled fractions (indicated P in larger chromatogram).

PCU IAS CENTRAL UNIVERSITY LIBRARY

14. Centrifuge the slightly cloudy solution 30 min at 22,000 × *g* (12,000 rpm in JA-14), 4°C. Decant the supernatant into a beaker and add an additional 17 g $(NH_4)_2SO_4$ per 100 ml solution (77% saturation or 3 M final concentration). Equilibrate with stirring and centrifuge 30 min at 22,000 × *g*, 4°C.

For the addition of $(NH_4)SO_4$ follow the same method as described in step 13.

15. Decant the supernatant and drain the pellets by inverting the tubes on a paper towel. Save the pellets.

Dialyze the fractionated sample

16. Suspend the pellets in ~300 ml cation-exchange buffer and dialyze, using Spectra/Por 1 dialysis tubing, against 5 liters cation-exchange buffer. Change the dialysis buffer at least once.

The dialysis step is conveniently performed overnight; the CM Sepharose column used in step 18 can be prepared during this period.

The dialysis tubing is prepared by heating 30 to 60 min at 90° to 95°C in 5 mM EDTA. The tubing is then washed well with water and stored in 10% ethanol at 4°C prior to use. A suitable length of tubing is filled to about one-half to three-quarters capacity with solution (to allow for expansion) and sealed with two knots at each end. Use gloves when handling the tubing, check for leaks before use, and make sure the magnetic stir-bar does not rub against the tubing. See APPENDIX 3B for further information concerning dialysis.

17. After dialysis, remove the slightly cloudy solution from the tubing and centrifuge 30 min at 22,000 × *g* (12,000 rpm in JA-14), 4°C. Save the supernatant.

Chromatograph dialyzed sample on cation-exchange resin

18. Prepare ~200 to 225 ml CM Sepharose CL-4B resin by washing on a sintered-glass funnel first with water, then with cation-exchange buffer (pH 5.7; wash as in step 1 except using different buffer). Pack the degassed resin into a 5 × 50–cm column as in step 2.

The packed column will have a bed height of ~11 to 12 cm. The comments made in the annotation to step 3 also apply here.

19. Elute the column using cation-exchange buffer at 100 to 150 ml/hr with a peristaltic pump. Monitor the column effluent at 280 or 260 nm using a suitable UV detector. Check that the pH and conductivity of the column effluent are the same as for the buffer applied to the column.

20. Check the pH and conductivity of the dialysate supernatant from step 17 and, if necessary, dilute with water so that the conductivity is in the range 1.0 to 1.2 mS/cm (at 4° to 6°C). Apply the clear solution to the CM Sepharose column at a flow rate of 150 ml/hr. When the UV absorbance of the column effluent approaches baseline, proceed to the next step.

21. Prepare a 0 to 250 mM NaCl gradient in cation-exchange buffer by adding 500 ml buffer to the inner chamber of the gradient maker and 500 ml buffer/250 mM NaCl to the outer chamber. Apply the gradient to the column at 150 ml/hr. Collect 15-ml fractions.

The total volume of the gradient is 1 liter (~4.5 column volumes).

At pH 5.7, IL-1β (pI 6.8) is positively charged and binds to the negatively charged cation-exchange resin. IL-1β is eluted from the column with ~100 mM NaCl, and it will be located in the major absorbance peak (see Wingfield et al., 1986, for figure of typical elution profile).

BOURNEMOUTH UNIVERSITY LIBRARY

22. Monitor the progress of the gradient using an in-line conductivity meter positioned after the absorbance flow cell. Assay column fractions for IL-1β by SDS-PAGE (*UNIT 10.1*) and pool fractions containing IL-1β.

Deciding what fractions to pool is dictated by the fact that the remaining purification stage is gel filtration, a method that will not remove contaminants with sizes close to that of IL-1β (17.4 kDa). As IL-1β is a well-expressed protein (>5% total protein), one can afford to be conservative and pool for purity rather than yield.

The used CM-Sepharose matrix can be cleaned up and stored as described in the annotation to step 12.

CM Sepharose Fast Flow or SP Sepharose FF (a strong cation exchanger; both resins are also from Pharmacia Biotech) can be used instead of CM Sepharose CL-4B. Similar results are obtained with either matrix, with the advantage of faster flow rates.

Concentrate the proteins

23a. *To concentrate proteins by ultrafiltration:* Adjust the CM Sepharose pool (150 to 250 ml) to pH 7.5 with Tris base. Assemble a 200- or 400-ml stirred ultrafiltration cell containing a washed ultrafilter membrane, either YM3 (3-kDa cutoff) or PM10 (10-kDa cutoff). Pressurize the cell with nitrogen according to manufacturer's recommendations. Collect the effluent from the cell into a measuring cylinder and occasionally record the absorbance at 280 nm to check that the membrane is not leaking. When the volume remaining in the cell is ~15 ml, depressurize the cell and carefully remove the solution using a Pasteur pipet with a small length of polyethylene tubing attached to the end so as not to scratch the membrane. Use ~3 ml gel-filtration buffer to wash the membrane, adding washings to the main concentrate.

Exposure to solutions below pH 4.5 and above pH 10.5 to 11.0 often causes protein denaturation. Hence, when adjusting the pH of protein solutions, there is less chance of overshooting the required pH by using concentrated buffer components instead of pure acid or base. For example, Tris base is used rather than dilute NaOH.

Select the ultrafiltration membrane pore size based on the size of the protein. The membrane can be reused; store in 5% ethanol at 4°C. It is good practice to reserve a membrane for use with a particular protein (compare comment on resin usage) and to wash the concentration cell carefully after use.

23b. *To concentrate proteins by salt precipitation:* Adjust the CM Sepharose pool to pH 7.5 with Tris base and slowly add 53.9 g $(NH_4)_2SO_4$ per 100 ml solution at 0°C (82.2% saturation or 3.2 M final concentration). Follow the basic salt precipitation method (steps 13 to 15) and suspend the pellets in 18 to 20 ml gel-filtration buffer.

Conduct gel-filtration chromatography

24. Prepare ~480 ml Ultrogel AcA54 resin by washing on a sintered-glass funnel with water, then gel-filtration buffer (as in step 1); suspend washed resin in gel-filtration buffer to 75% settled gel/25% buffer by volume and degas (step 2). Pour a slurry of degassed resin into a 2.5 × 100–cm chromatography column fitted with a filling reservoir. Pack the column at ~35 ml/hr.

UNITS 6.3 & 8.3 should be consulted for further details on the preparation and elution of gel-filtration columns. The column should be prepared in advance so that the protein samples can be applied as soon as the concentration step is completed.

The resin should be free-flowing yet concentrated enough to produce a packed bed with one pouring. A freshly packed gel-filtration column can be checked for packing irregularities by prerunning the column with a few colored markers. Blue dextran will be excluded from the gel matrix and will elute at the void volume (V_0, which equals 30% to 35% of the total column volume); cytochrome c (red; 12.4 kDa) will elute close

PCU IASI CENTRAL UNIVERSITY LIBRARY

to the expected position of IL-1β; and potassium dichromate (yellow) will be fully included in the gel matrix and will elute at the included volume (V_i ~480 ml).

Superdex 75 gel-filtration matrix (Pharmacia Biotech) can be substituted for Ultrogel AcA54. The former allows higher flow rates and thus is more compatible with the FPLC and BioPilot systems. Both Superdex and the cation-exchange Fast Flow resins mentioned above can be purchased in various prepacked columns that are useful for method development.

25. Filter the concentrated protein using a Millex-GV 0.22-μm filter unit attached to a 10- or 20-ml syringe and apply to the gel-filtration column. The sample can be applied either directly to the top of the column using a Pasteur pipet (care is required to prevent breaking the tip and contaminating the column) or via a three-way valve and syringe without removing the top flow adapter.

The volume of sample applied to the column (20 ml) represents ~4% of the total column volume. For columns with different dimensions, apply the same proportionate volume of sample. The column should be ≥60 cm long. For analytical separations, the sample volume should not exceed 2.5% of the column volume.

26. Elute the column at 35 ml/hr with gel-filtration buffer and collect 10-ml fractions. The major eluting peak contains the IL-1β; monitor the fractions by SDS-PAGE (*UNIT 10.1*) and pool fractions that contain pure protein. Save side fractions that contain small amounts of contaminants; this material can be rechromatographed after concentration as described in step 23a or 23b.

See Figure 6.2.1A for an example of results from SDS-PAGE.

Once the sample has been run into the column, it can be eluted with any buffer or even with a suitable column storage solvent such as 5 mM sodium azide in water.

Concentrate and store purified protein

27. Concentrate the purified protein, if required, by ultrafiltration (see step 23a). Determine the protein concentration of purified IL-1β by measuring absorbance at 280 nm.

A molar absorbance coefficient (ε) of 10.61 mM^{-1} cm^{-1} is used. This corresponds to an absorbance of 0.63 for a 1 mg/ml solution using a 1-cm-path length cell.

28. For short-term storage (≤12 months), filter the protein with a Millex-GV 0.22-μm filter unit, divide the solution into sterile plastic vials, and freeze aliquots rapidly with dry ice/ethanol. Store at –80°C. For long-term storage of IL-1β (>12 months), lyophilize the protein. Dialyze the sample using a volatile buffer such as 50 mM ammonium bicarbonate or a nonvolatile buffer such as lyophilization buffer.

To circumvent the dialysis step, the phosphate-based lyophilization buffer can be used for gel filtration instead of the Tris·Cl gel-filtration buffer.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX. Buffer pH and conductivities are for solutions at 4° to 6°C. Units of conductance are given in siemens ($S = \Omega^{-1}$).

***Anion-exchange buffer** (50 mM Tris·Cl, pH 8.5)*

Dilute 1 M Tris·Cl, pH 8.0 20-fold with water and adjust to pH 8.5 with NaOH. Make immediately before use.

Conductivity of the solution is 1.57 mS/cm.

B. C. I. IAȘI

BCU IAȘI / CENTRALA UNIVERSITARĂ LIBRARY

Cation-exchange buffer, 10× (15 mM sodium phosphate, pH 5.7/1 mM sodium azide)

19.4 g $NaH_2PO_4 \cdot H_2O$
1.4 g $Na_2HPO_4 \cdot 2H_2O$
0.65 g sodium azide
H_2O to 1 liter
Store up to 1 month at 4°C
Dilute 10-fold immediately prior to use

Conductivity of the solution is 0.9 mS/cm. Sodium azide is an antibacterial agent.

Cation-exchange buffer (1×)/250 mM NaCl

14.61g NaCl
100 ml 10× cation-exchange buffer (see recipe)
H_2O to 1 liter

***Gel-filtration buffer** (100 mM Tris·Cl, pH 7.5/1 mM sodium azide)*

Dilute 1 M Tris·Cl, pH 8.0 10-fold with water and adjust to pH 7.5 with HCl. Add sodium azide from a 1 M stock solution (65 g/liter). Make immediately before use.

***Lyophilization buffer** (25 mM sodium phosphate, pH 7.5/0.5 mM sodium azide)*

2.2 g $NaH_2PO_4 \cdot H_2O$
11.9 g $Na_2HPO_4 \cdot 2H_2O$
1.3 g sodium azide
H_2O to 4 liters
Make fresh and use immediately

Usually, 2 liters is required for dialysis, with one buffer change (4 liters total). Gel filtration requires smaller volumes.

Lysis buffer

100 mM Tris·Cl, pH 8.0
2 mM EDTA, pH 8.0
5 mM benzamidine·HCl (780 mg/liter)

Make immediately prior to use; alternatively, make ahead of time and store up to several days at 4°C.

Conductivity of the solution is 1.57 mS/cm.

It should be noted that a 1°C decrease in temperature increases the pH of the Tris buffer by ~0.03 pH units. Both Tris·Cl and EDTA stock solutions are commercially available (e.g., Life Technologies).

Benzamidine·HCl is a water-soluble serine protease inhibitor. An alternative is 50 μM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF; Perfabloc SC, Boehringer Mannheim), a water-soluble inhibitor with the same spectrum of activity as phenylmethylsulfonyl fluoride (PMSF).

COMMENTARY

Background Information

Expression of soluble proteins

The basic stages involved in preparing soluble proteins from *E. coli* are achieving protein expression, harvesting and breaking cells (or collecting culture medium, in the case of secreted proteins), purifying proteins, and characterizing the purified protein. Choices that must be made at each stage are briefly reviewed.

Optimizing the expression system

The basic requirements for protein expression in *E. coli* are discussed in UNIT 5.1. Experience has shown that it is usually possible to express any given protein in *E. coli*, but whether the protein will be expressed in a soluble and unmodified state is unpredictable. Tinkering with the various elements of the expression plasmid, changing the *E. coli* host strains, and altering the fermentation conditions are carried out to optimize the expression

PCU IASI / CENTRAL UNIVERSITY LIBRARY

of unmodified soluble protein (*UNIT 5.1*). Use of secretion vectors and fusion protein constructs to prevent aggregation has been discussed elsewhere (*UNITS 5.1 & 6.1*). Even if efforts to express soluble protein fail and inclusion bodies are produced, there is often a good chance that the protein can be extracted and folded into native protein (see *UNITS 6.4 & 6.5*).

There are situations where the intrinsic properties of a protein can preclude its expression in a soluble and stable form. Examples include the following: (1) individual subunits of a heterodimer, which may be unstable and may degrade or aggregate when expressed separately (coexpression of the subunits may solve the problem; Nash et al., 1987); (2) proteins derived from activated precursors (*UNIT 6.4*; Hlodan and Hartl, 1994); and (3) proteins in which the native (authentic) sequences have been modified by site-directed mutagenesis (Chrunyk et al., 1993). It is worth noting that the solubility of proteins that are otherwise insoluble or poorly soluble can, conversely, sometimes be improved by site-directed mutagenesis (Dyda et al., 1994).

Once conditions have been found that result in the expression of soluble protein, the next hurdle is to recover the protein in an unmodified form. Avoiding proteolysis and chemical modification during the isolation process is crucial.

Breaking cells

Methods for breaking cells or selectively extracting proteins are reviewed in *UNIT 6.1*. In the protocol for purifying human IL-1β, cells are disrupted with a French press. This method is the most efficient way to break *E. coli* cells and is the method of choice. The equipment, however, requires dedicated laboratory space and may be expensive for some laboratories. Cell breakage with lysozyme (*UNIT 6.5*) and/or sonication is also frequently used; such methods are especially suited to small-scale work (reviewed by Hopkins, 1991).

It is important to consider whether the cell breakage method influences the final solubility and stability of the recombinant protein in the cell lysates. A soluble recombinant protein constituting ~12% of the total bacterial protein will be present in the cytoplasmic space at a concentration of at least 30 mg/ml, assuming a total protein and RNA concentration in the cytoplasm of ~340 mg/ml in nontransformed cells (Zimmerman and Trach, 1991) and a 2.7:1 weight ratio of protein to RNA (Neidhart, 1987). Cell breakage can be expected to increase the solubility of the recombinant protein by a dilution effect (discussed by Zimmerman and Trach, 1991). Cell lysis procedures, however, especially when mechanical shear is used (e.g., French press and sonication), can produce local heating that may have a denaturing effect and lead to aggregation. Mechanical shearing will also be expected to decrease the size of nucleic acids and release polyanionic lipopolysaccharides; both classes of coumpounds can bind nonspecifically to proteins, especially basic proteins, thereby reducing their solubility. Including detergents in the lysozyme treatment may also release potent membrane-bound proteases.

Despite these potential pitfalls, the recovery of soluble protein is usually not dramatically influenced by the cell breakage method, but there is always the potential that it might be. Protein recovery as a function of cell breakage methodology is rarely systematically studied or reported. Johnson and Hecht (1994) reported that freezing and thawing of cells selectively releases recombinant proteins located in the cytoplasm. The freeze-thaw procedure, in principle, should be the best method for releasing soluble proteins in an unmodified state, but its general applicability remains to be seen.

Determining solubility

The solubility of recombinant protein in the cell lysate or extract is usually established by differential centrifugation (*UNITS 6.1 & 5.3*). Alternatively, the extract can be clarified by filtering and applied to a small gel-filtration column. An HPLC or FPLC gel-filtration separation in conjunction with a rapid screening method for assaying column fractions (e.g., SDS-PAGE using PhastSystem from Hoefer Pharmacia) can produce a determination of both solubility and size distribution in ~2 hr or less.

Purifying protein

There are many approaches to purifying proteins, such as expressing fusion proteins that contain engineered affinity handles or tags (see *UNITS 5.1 & 6.1* for further details). Most soluble proteins that are well behaved—i.e., do not contain excessive charge heterogeneity—are non-self-associating and can often be purified by two stages of ion-exchange chromatography and one or more polishing steps, one of which should be gel filtration. The basic approach is outlined in Figure 6.1.3. The purification of human IL-1β is a typical example. Specific accounts of the various chromatographic methods are found in Chapters 8 and

BOULDER LIBRARY UNIVERSITY

9; Scopes (1993) particularly emphasizes first principles.

Characterizing protein

Once a protein has been purified, it is usually characterized (Chapter 7) in order to establish chemical homogeneity (primary sequence) and physical homogeneity (size and conformation). Not only is this information required for a fundamental description of the protein, but it will also give insights on how best to rationalize the purification process. Moreover, the detection of certain heterogeneities may give clues on how to avoid their occurrence. For example, chemical modifications such as deamidation may be prevented by changing one or more purification steps that expose the protein to extremes of pH. Other modifications that arise from construction of the expression plasmid itself, once recognized, can often be fixed.

Characteristics of recombinant and authentic interleukin 1β

The cytokine IL-1β and the closely related IL-1α are synthesized mainly by monocytes and mononuclear cells as 31-kDa cytoplasmic precursor proteins (reviewed by Dinarello, 1989). A unique cytoplasmic cysteine protease (IL-1β-converting enzyme, or ICE) cleaves the IL-1β precursor in half, generating the mature form of the protein comprising the C-terminal 153 residues. The biological activities of IL-β (and IL-1α) are initiated by interaction with either type I or type II cellular receptors (McMahan et al., 1991). Both receptors have three immunoglobulin-like extracellular ligand-binding domains and single membrane-spanning segments. Another form of IL-1, the interleukin 1 receptor antagonist (IL-1ra), is a natural competitive antagonist of IL-1β.

Authentic IL-1β been isolated from human monocyte cell cultures by HPLC using anion-exchange and gel-filtration (or reversed-phase) matrices. A 3000-fold purification was required to obtain 16 μg protein from 5 liters of cell culture extract (Gery and Schmidt, 1985). The authentic protein was established to be a non-glycosylated monomer (~18 kDa), with a pI of 6.8 and containing a single N-terminal sequence (AlaProValArg) as predicted by the DNA coding sequence.

The recombinant protein produced in *E. coli* has the same properties as the natural product except that N-terminal heterogeneity is often observed. The recombinant protein usually contains a mixture of forms with Met (20%), Ala (67%), or Pro (13%) as the N terminus. The Met form is derived by incomplete processing of the initiating Met, and the Pro form is due to cleavage of the authentic N-terminal Ala by *E. coli* protease. Despite the slight differences among the recombinant variants, partial or complete separations have been achieved (Wingfield et al., 1987; Yem et al., 1988).

Both the crystal structure of IL-1β (Priestle et al., 1988) and its nuclear magnetic resonance (NMR) structure in solution (Clore et al., 1991) have been determined using material purified by the protocol described here. For some NMR studies it was necessary to produce material with a single N terminus, and this was fortuitously achieved by replacing the N-terminal Ala residue with a Cys residue. Purified protein from cells expressing the mutant (N-terminal CysProValArg) contain only N-terminal Pro. It should be noted that deletion of the first four residues of IL-1β has no effect on activity; however, the presence of the unprocessed Met reduces receptor binding 10-fold.

Interleukin 1β does not contain a disulfide—the cysteines are either buried (Cys-71) or partially buried (Cys-8)—so there is no need to include reductants in column buffers. As mentioned earlier, the authentic protein is non-glycosylated even though there is a potential N-glycosylation site (-Asn^{7}-Cys^{8}-Thr^{9}-) near the N terminus. Interestingly, when IL-1β is expressed in yeast, equal amounts of N-glycosylated (21 kDa) and nonglycosylated protein (17 kDa) are produced (Livi et al., 1991).

Other approaches to purifying recombinant interleukin 1β

Because of the biomedical importance of IL-1β, several reports have described expression and purification of the protein. Some examples are discussed.

Kronheim et al. (1986) acidified *E. coli* cell lysates to pH 4 and thus precipitated ~70% of the total protein while most of the IL-1β remained soluble. IL-1β was further purified by cation-exchange, anion-exchange, and dye-matrix chromatographies (*UNITS 8.2 & 9.2*); ~200 mg protein was recovered from 2.5 liters of culture. The noteworthy feature of this method is the acid precipitation stage. Extracts of *E. coli* contain predominately negatively charged (anionic) proteins: 60% have pI values between 5.0 and 6.0 and 80% between 4.5 and 6.7 (Sherwood, 1992). Adjustments of cell extracts to pH 5.0 and below thus result in isoelectric precipitation of substantial amounts of *E. coli* protein. This is a useful purification step as long

PCU IASRI CENTRAL UNIVERSITY LIBRARY

as the recombinant protein remains soluble (or does not coprecipitate) under the conditions.

In another approach, Meyers et al. (1987) carried out ammonium sulfate fractionation directly on the *E. coli* extract. The IL-1β protein was recovered in the fraction corresponding to 50% to 80% saturation. After desalting, the protein was further purified by ion-exchange (cation and anion) and gel-filtration chromatographies; ~400 mg protein was recovered from 250 g wet weight cells.

The high-speed centrifugation step used to remove particulate material (see Basic Protocol, step 6) can often be replaced by salt fractionation. For IL-1β, advantage can be taken of the fact that a relatively high $(NH_4)_2SO_4$ concentration is required to precipitate the protein. The disadvantage of using this approach early in the purification is that the protein must be desalted before being applied to an ion-exchange column.

Finally, it is of interest to note that IL-1β can be released from *E. coli* cells by osmotic shock treatment (Joseph-Liauzun et al., 1990). It is well known that certain *E. coli* cytoplasmic proteins, including thioredoxin, can be released by osmotic shock.

Scale of procedure

The Basic Protocol described starts with 50 g cell paste and yields ~400 mg pure IL-1β. For smaller (1 to 5 g) or larger (100 g) cell quantities, the method can be adapted by simply reducing or increasing, respectively, the cross-sectional area of the chromatography columns. For example, for 5 g cells, the diameters of the DEAE and CM Sepharose CL-4B columns are reduced to 2.5 cm, with column lengths of 20 and 10 cm, respectively. This results in 1:4 reductions in the resin volumes compared to those for the protocol utilizing 50 g cells. The original gel-filtration column can be used (2.5 × 100 cm); however, the volume of ~1:4 sample applied should be reduced to 8 to 10 ml. Alternately, a column of about the same length (60 to 100 cm) but smaller diameter (e.g., 1.25 cm) should be used. For discussion on scaling up column chromatography procedures, see *UNIT 8.3* (Strategic Planning and Scopes, 1993).

Critical Parameters and Troubleshooting

High-level protein expression with good cell yield (*UNITS 5.1-5.3*) is required to obtain reasonable amounts of pure protein. This is obviously true no matter how efficient (or inefficient) the purification protocol. Once the host/vector system and the fermentation conditions have been optimized and, equally important, standardized, the level and solubility of recombinant protein in the starting material should be consistent. However, it is sensible to check both the expression level and solubility before starting each protein purification. The expression level can usually be monitored by SDS-PAGE; *UNIT 5.2* describes how to prepare samples for analysis by SDS-PAGE, and *UNIT 10.1* gives electrophoresis procedures. Solubility can be easily monitored by breaking a small amount of cells (≤1.0 g) by sonication (*UNIT 5.3*) or by treating cells with lysozyme (*UNIT 6.5*) and microcentrifuging the resulting cell lysate to judge whether the protein is highly aggregated.

There are many reasons why protein expression may not be as high as anticipated and, worse, why the expressed protein may switch from being soluble to insoluble. For example, faulty pH control during the fermentation may cause insolubility. Whatever the situation, one should be in a position to rationally decide whether to continue with the purification or repeat the initial fermentation.

If the protein expression level is normal (>5%) and the protein has accumulated in a soluble state, implementing the Basic Protocol, which uses only standard purification methods, should be trouble-free. Common mistakes involve buffer preparation and other trivial errors such as pooling the wrong column fractions. As discussed later, do not discard column fractions until they have been completely analyzed.

A classic purification table (see for example, Table III.1 in Dixon and Webb, 1979), has not been presented, as the biological activities of IL-1β could not be measured reliably in crude *E. coli* cell extracts by assays available to the author when the method was developed (Wingfield et al., 1986). As is often the case with proteins requiring complex biological assays, protein size analysis by SDS-PAGE is used to follow purification. During the initial stages, it is of course necessary to confirm the identity of the particular band one is following, and this can be accomplished by eluting or blotting the protein from SDS-polyacrylamide gels and performing N-terminal amino acid sequence analysis. Immunoblotting with specific antibody can be used to monitor protein purification when the recombinant protein is in low abundance (e.g., for secreted proteins in the cell medium). Finally, it should be noted that several manufacturers (e.g., R&D Systems and Genzyme) now supply assay kits for many cytokines, including IL-1β.

ROSWELL PARK CANCER INSTITUTE LIBRARY

Cell breakage

Efficient cell breakage should be trouble-free as long as the French press is operated in accordance with the manufacturer's instructions. It takes a little practice to operate the flow valve. The aim is to generate as high a flow rate as possible while maintaining a pressure gauge reading of ~1000. If the flow rate is too fast, the pressure reading will drop and unbroken cells will pass into the flow stream. Toward the end of the run, the flow rate should be reduced, as it becomes difficult to control the pressure. After use the French pressure cell should be cleaned and dried, and the flow valve ball should be replaced. Store the cell at 4°C.

Protein purification

In general, troubleshooting a purification method will be much easier if fractions are not discarded until the appropriate monitoring of the purification steps is complete. When using an established method, fractions from a chromatographic run are frequently pooled on the basis of absorbance only and the remainder quickly discarded. Ammonium sulfate supernatants or pellets are often discarded on the basis of previous fractionation behavior or pilot-scale work. As stated, do not discard any fractions until they have been checked, usually by SDS-PAGE. If there has been a problem, it can usually be easily sorted out if all the fractions from the various stages are still available. If in doubt about conditions for storing fractions, freeze selected fractions at as low a temperature as possible (ideally –80°C) and discard when appropriate. When freezing material, it is worthwhile to take the extra effort to dispense small samples (<1 ml) into microcentrifuge tubes amenable for rapid analysis, if necessary, to avoid having to thaw large volumes merely to take 10 µl for SDS-PAGE.

For reproducible results, careful recording of the pH and conductivity of starting buffers, protein solutions, and column eluents is necessary. For example, if the protein does not bind to the cation exchanger (step 19), the ionic strength of the sample or column buffer may be too high. Alternatively, the buffer pH may be too high. Errors of this kind are avoided by simply measuring the conductivity and pH of all solutions at each stage. Other precautions and critical steps relating to chromatography procedures are mentioned in Chapters 8 and 9.

Anticipated Results

Purification results in recovery of ~8 mg purified protein per gram wet weight cells. That result is based on a 10% expression level; IL-1β expression levels of 10% to 15% are routinely obtained with the expression system previously described (Wingfield et al., 1986). With a 1.5-liter fermenter, at least 50 g wet weight cells/liter can be obtained using a standard complex medium (*UNIT 5.3*). Thus, ~400 mg protein is obtained from 1 liter of culture, and optimal conditions can lead to nearly twice that amount. The yield corresponds to an overall recovery of ~60% of the IL-1β originally in the cells.

Figure 6.2.1 illustrates the protein content at various purification stages. The starting material for the final stage of gel filtration (Fig. 6.2.1A, lane d) is mainly contaminated with *E. coli* proteins >44 kDa in size, and these are easily separated from the IL-1β protein (17.4 kDa) by gel filtration. Analytical rechromatography of the protein from the pooled fractions (indicated P in Fig. 6.2.1B) reveals a single symmetrical peak indicative of purity and physical homogeneity (Fig. 6.2.1B inset). A typical cation-exchange chromatogram for the purification appears elsewhere (Wingfield et al., 1986).

Time Considerations

Because IL-1β appears to be stable against proteolytic degradation and other chemical modications during purification, the speed of purification was not critical in this case. The low-pressure chromatographic method described in the Basic Protocol requires ~4 days, which can be shortened to ~3 days using the Pharmacia Biotech BioPilot or FPLC systems in conjunction with matrices that allow faster flow rates. The times required for purification whether using low-pressure chromatography (as described in the Basic Protocol) or medium-pressure chromatography in the FPLC or BioPilot systems are summarized in Table 6.2.1.

Literature Cited

Burgess, R.R. and Jendrisak, J.J. 1975. A procedure for the rapid, large-scale purification of *E. coli* DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *J. Biol. Chem.* 14:4634-4638.

Chrunyk, B.A., Evans, J., Lillquist, J., Young, P., and Wetzel, R. 1993. Inclusion body formation and protein stability in sequence variants of Interleukin-1β. *J. Biol. Chem.* 268:18053-18061.

Clore, G.M., Wingfield, P.T., and Gronenborn, A.M. 1991. High-resolution structure of interleukin 1β in solution by three- and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry* 30:2315-2323.

PCU IASI CENTRAL UNIVERSITY LIBRARY

Dinarello, C.A. 1989. Interleukin-1 and its biologically related cytokines. *Adv. Immunol.* 44:153-205.

Dixon, M. and Webb, E. 1979. Enzyme isolation. *In* Enzymes (3rd ed.) pp. 23-46. Academic Press, New York.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., and Davies, D.R. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: Similarity to other polynucleotidyltransferases. *Science* 266:1981-1986.

Gery, I. and Schmidt, J.A. 1985. Human interleukin 1. *Methods Enzymol.* 116:456-467.

Hlodan, R. and Hartl, F.U. 1994. How the protein folds in the cell. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 194-228. IRL Press, Oxford.

Hopkins, T.R. 1991. Physical and chemical cell disruption for the recovery of intracellular proteins. *In* Purification and Analysis of Recombinant Proteins (R. Seetharam and S.K. Sharma, eds.) pp. 57-83. Marcel Dekker, New York.

Johnson, B.H. and Hecht, M.H. (1994) Recombinant proteins can be isolated from *E. coli* by repeated cycles of freezing and thawing. *Bio/Technology* 12:1357-1360.

Joseph-Liauzun, E., Legoux, R., Guerveno, V., Marchese, E., and Ferra, P. 1990. Human recombinant interleukin-1β isolated from *E. coli* by simple osmotic shock. *Gene* 86:291-295.

Kronheim, S.R., Cantrell, M.A., Deeley, M.C., March, C.J., Glackin, P.J., Anderson, D.M., Hemenway, T., Merriam, J.E., Cosman, D., and Hopp, T.P. 1986. Purification and characterization of human interleukin-1 expressed in *Escherichia coli*. *Bio/Technology* 4:1078-1082.

Livi, G.P., Lillquist, J.S., Ferrara, A., Sathe, G.M., Simon, P.L., Meyers, C.A., Gorman, J.A., and Young, P.R. 1991. Secretion of N-glycosylated interleukin-1β in *Saccharomyces cerevisiae* using a leader peptide from *Candida albicans*. Effect of N-linked glycosylation on biological activity. *J. Biol. Chem.* 266:15348-15348.

McMahan, C.J., Slack, J.L., Mosley, B., Cosman, D., Lupton, S.D., Brunton, L.L., Grubin, C.E., Wignall, J.M., Jenkins, N.A., Brannan, C.I., Copeland, N.G., Huebner, K., Croce, C.M., Cannizzarro, L.A., Benjamin, D., Dower, S.K., Spriggs, M.K., and Sims, J.E. 1991. A novel IL-1 receptor, cloned from B cells by mammalian expression, is expressed in many cell types. *EMBO J.* 10:2821-2832.

Meyers, C.A., Johanson, K.O., Miles, L.M., McDevitt, P.J., Simon, P.L., Webb, R.L., Chen, M.-J., Holskin, B.P., Lillquist, J.S., and Young, P.R. 1987. Purification and characterization of human recombinant interleukin-1β. *J. Biol. Chem.* 262:11176-11181.

Nash, H.A., Robertson, C.A., Flamm, E., Weisberg, R.A., and Miller, H. 1987. Overproduction of *Escherichia coli* integration host factor, a protein with nonidentical subunits. *J. Bacteriol.* 169:4124-4127.

Neidhardt, F.C. 1987. Chemical composition of *Escherichia coli*. *In Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology (F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger, eds.) pp. 3-6. American Society for Microbiology, Washington, D.C.

Priestle, J.P., Schar, H.-P., and Grutter, M.G. 1988. Crystal structure of the cytokine interleukin 1β. *EMBO J.* 7:339-343.

Scopes, R.K. 1994. Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.

Sherwood, R.F. 1992. Making bacterial extracts suitable for chromatography. *Meth. Mol. Biol.* 11:287-305.

Wingfield, P., Payton, M., Tavernier, J., Barnes, M., Shaw, A., Rose, K., Simona, M.G., Demczuk, S., Williamson, K., and Dayer, J.M. 1986. Purification and characterization of human interleukin-1β expressed in recombinant *Escherichia coli*. *Eur. J. Biochem.* 160:491-497.

Wingfield, P.T., Graber, P., Rose, K., Simona, M.G., and Hughes, G.J. 1987. Chromatofocusing on N-terminally processed forms of proteins. *J. Chromatogr.* 387:291-300.

Wood, W.I. 1976. Tables for the preparation of ammonium sulfate solutions. *Anal. Biochem.* 73:250-257.

Yem, A.W., Richard, K.A., Staite, N.D., and Deibel, M.R. 1988. Resolution and biological properties of three N-terminal analogues of recombinant human interleukin-1β. *Lymphokine Res.* 7:85-92.

Zimmerman, S.B. and Trach, S. 1991. Estimation of macromolecular concentrations and excluded volume effects in the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* 222:599-620.

Key Reference

Wingfield et al., 1986. See above.

The original publication on which Basic Protocol 1 is based.

Contributed by Paul T. Wingfield
National Institutes of Health
Bethesda, Maryland

BCU Cardiff Central University Library

Preparation and Extraction of Insoluble (Inclusion-Body) Proteins from *Escherichia coli*

High-level expression of many recombinant proteins in *Escherichia coli* leads to the formation of highly aggregated protein commonly referred to as inclusion bodies (*UNITS 5.1 & 6.1*). Inclusion bodies are normally formed in the cytoplasm; alternatively, if a secretion vector is used, they can form in the periplasmic space.

Inclusion bodies recovered from cell lysates by low-speed centrifugation are heavily contaminated with *E. coli* cell wall and outer membrane components. The latter are largely removed by selective extraction with detergents and low concentrations of either urea or guanidine·HCl to produce so-called washed pellets. These basic steps result in a significant purification of the recombinant protein, which usually makes up ~60% of the washed pellet protein. The challenge, therefore, is not to purify the recombinant-derived protein, but to solubilize it and then fold it into native and biologically active protein.

Basic Protocol 1 describes preparation of washed pellets and solubilization of the protein using guanidine·HCl. The extracted protein, which is unfolded, is either directly folded as described in *UNIT 6.5* or further purified by gel filtration in the presence of guanidine·HCl as in Basic Protocol 2. A support protocol describes the removal of guanidine·HCl from column fractions so they can be monitored by SDS-PAGE (*UNIT 10.1*).

PREPARATION AND EXTRACTION OF INSOLUBLE (INCLUSION-BODY) PROTEINS FROM *ESCHERICHIA COLI*

BASIC PROTOCOL 1

Bacterial cells are lysed using a French press, and inclusion bodies in the cell lysate are pelleted by low-speed centrifugation. The pellet fraction is washed (preextracted) with urea and Triton X-100 to remove *E. coli* membrane and cell wall material. Guanidine·HCl (8 M) and dithiothreitol (DTT) are used to solubilize the washed pellet protein. Extraction with the denaturant simultaneously dissociates protein-protein interactions and unfolds the protein. As a result, the extracted protein consists (ideally) of unfolded monomers, with sulfhydryl groups (if present) in the reduced state.

Materials

E. coli cells from fermentation (*UNIT 5.3*) containing the protein of interest
Lysis buffer (see recipe)
Wash buffer (see recipe), with and without urea and Triton X-100
Extraction buffer (see recipe)

250- and 500-ml stainless steel beakers
0.22-μm syringe filters (e.g., Millex from Millipore)
20-ml disposable syringe

Additional equipment for breaking cells, homogenizing cells and pellets and centrifuging at low and high speeds (*UNIT 6.2*)

Break cells and prepare clarified lysate

1. Place thawed *E. coli* cells in a stainless steel beaker. Add 4 ml lysis buffer per gram wet weight of cells. Keep bacterial cells cool by placing the beaker on ice in an ice bucket.

The cells can be pretreated with lysozyme prior to lysis in the French press. Lysozyme treatment involves incubating cells ~20 min at 20° to 25°C in lysis buffer supplemented

Copyright © 1995 by John Wiley & Sons, Inc.

BOWLING GREEN STATE UNIVERSITY LIBRARY

with 200 μg/ml lysozyme, with intermittent homogenization using a tissue grinder. It should be emphasized that this optional step is carried out before French press breakage and is not simply an alternative method of cell breakage (compare the comments made in the annotation to step 4 of UNIT 6.2). Its purpose is to aid removal of the peptidoglycan and outer membrane protein contaminants during the washing steps (steps 6 to 9; for further details see UNIT 6.1 and Fig. 6.1.5). An example of this approach is given in Basic Protocol 1 of UNIT 6.5.

2. Suspend cells using a Waring blender and homogenize using the Polytron tissue-grinder homogenizer until all clumps are disrupted, as described in *UNIT 6.2*, step 3.

3. Lyse cells with two passes through the French pressure cell operated at 16,000 to 18,000 lb/in^2 (with the high-ratio setting, pressure gauge readings between 1011 and 1135), chilling the cell suspension to 4°C after each pass, as described in *UNIT 6.2*, steps 2 and 4.

4. Reduce the viscosity of the suspension by sonicating 5 min at full power with 50% duty cycle (on for 5 sec, off for 5 sec) using an ultrasonic homogenizer, as described in *UNIT 6.2*, step 5.

5. Clarify the lysed cell suspension by centrifuging 1 hr at 22,000 × *g* (12,000 rpm in a JA-14 rotor in a Beckman J2-21M centrifuge), 4°C.

   Unbroken cells, large cellular debris, and the inclusion body protein will be pelleted.

   The JA-14 rotor uses 250-ml centrifuge bottles. For processing smaller volumes the Beckman JA-20 rotor (or equivalent) with 50-ml tubes can be used, at 13,500 rpm (22,000 × g).

   The procedure for dealing with insoluble inclusion-body proteins now diverges from that for purifying soluble proteins (UNIT 6.2).

Prepare washed pellets

6. Carefully pour off the supernatant from the pellet. Using a tissue homogenizer, suspend the pellet with 4 to 6 ml wash buffer per gram wet weight cells.

   Complete homogenization of the pellet is important to wash out soluble proteins and cellular components. Removal of cell wall and outer membrane material can be improved by increasing the amount of wash solution to 10 ml per gram cells.

   The concentration of urea and Triton X-100 in the wash buffer can be varied. The urea concentration is usually between 1 and 4 M; higher concentrations may result in partial solubilization of the recombinant proteins. The usual detergent concentration is 0.5% to 5%. Triton X-100 will not solubilize inclusion body proteins; it is included to help extract lipid and membrane-associated proteins.

7. Centrifuge the suspension 30 min at 22,000 × *g* (12,000 rpm in JA-14), 4°C. Discard supernatant and, using the tissue homogenizer, suspend the pellet in 4 to 6 ml wash buffer per gram wet weight of cells.

8. Repeat step 7 two more times.

   If the supernatant is still cloudy or colored, continue washing the pellet until the supernatant is clear.

9. Suspend the pellet with wash buffer minus the Triton X-100 and urea, using 4 to 6 ml buffer per gram wet cells. Centrifuge 30 min at 22,000 × *g* (12,000 rpm in JA-14), 4°C.

   The final wash removes excess Triton X-100 from the pellet.

PCU IASI - CENTRAL UNIVERSITY LIBRARY

If necessary the washed pellets can be stored at –80°C. It is better to store material at this stage rather than after the extraction stage (see comments to step 13).

Extract recombinant protein from washed pellets with guanidine·HCl

10. Using the tissue homogenizer, suspend the pellet with guanidine·HCl-containing extraction buffer. Use 0.5 to 1.0 ml buffer per gram wet weight of original cells if the extract will be subjected to gel filtration, and 2 to 4 ml buffer if the extract will be used in protein folding procedures. Perform this step at room temperature.

    To estimate the amount of recombinant protein in the washed pellets, use the following guidelines. (1) An expression level of 1% corresponds to ~1 mg recombinant protein per 1 g wet cells. (2) The recovery of highly aggregated recombinant protein in the washed pellets is ~75% that originally present in the cells. (3) About 60% of the total washed pellet protein is recombinant-derived. Thus, if 50 g cells is processed and the expression level is 5%, the washed pellets contain ~200 mg recombinant protein.

    The total amount of recombinant-derived protein in washed pellets can be directly determined by measuring the total protein concentration or by analyzing the washed pellets via SDS-PAGE (see Support Protocol and UNIT 10.1) to determine the proportions of the protein constituents.

    For gel-filtration purposes, the pellets from 50 g wet weight E. coli cells are solubilized with 40 to 50 ml extraction buffer (see Basic Protocol 2); the concentration of recombinant protein in the extract will be 4 to 5 mg/ml. For direct protein folding (UNIT 6.5), the pellets are extracted with 100 to 200 ml buffer, and the concentration of recombinant protein 1 to 2 mg/ml. If the washed pellet is heavily contaminated with outer cell wall and peptidoglycan material, the extract must be diluted further with extraction buffer (usually 1:1 to 1:3) to reduce the viscosity before it can be used for chromatography.

11. Centrifuge the suspension 1 hr at 100,000 × *g* (30,000 rpm in Ti45 rotor in a Beckman Optima XL-90 ultracentrifuge), 4°C.

    For volumes <250 ml the Beckman 70Ti rotor (capacity 6 × 39 ml) can be used at 32,000 rpm (~100,000 × g).

12. Carefully pour off the supernatant from the pellet. Filter the supernatant through a 0.22-µm syringe filter attached to a 20-ml disposable syringe.

    The filter removes unpelleted large cell wall debris that will clog most chromatography columns.

13. Use the clarified inclusion body extract for preparing folded protein (*UNIT 6.5*) or purify further by gel filtration (see Basic Protocol 2).

    The extract can be stored at –80°C until required. Freeze in plastic (or polyethylene) containers rather than glass. Divide sample into 10- to 20-ml aliquots instead of freezing in one large lot and fill containers to only 50% to 75% capacity.

MEDIUM-PRESSURE GEL-FILTRATION CHROMATOGRAPHY IN THE PRESENCE OF GUANIDINE HYDROCHLORIDE

BASIC PROTOCOL 2

Washed, extracted pellets (see Basic Protocol 1) contain >50% recombinant protein and are used as the starting material for purification of the protein of interest by gel-filtration chromatography. Superdex 200 gel-filtration medium, which allows high flow rates, is washed and packed into a column. The column is equilibrated at 4°C and the sample is applied.

Assay of column fractions by gel electrophoresis in the presence of SDS is complicated by the fact that guanidine·HCl forms a precipitate with SDS. Therefore, preparing samples

BOURNS / CENTRAL UNIVERSITY LIBRARY

for gel analysis involves selective precipitation of protein from guanidine·HCl prior to SDS-PAGE (see Support Protocol). The purified (or partially) purified protein is used as the starting material for procedures (e.g., *UNIT 6.5*) in which the denatured protein is folded into a native and biologically active structure.

Materials

Gel-filtration medium: Superdex 200 PG (preparative grade; Pharmacia Biotech)
5% (v/v) ethanol
Gel-filtration buffer (see recipe)
Guanidine·HCl extract of *E. coli* cells containing the protein of interest (see Basic Protocol 1)

4- to 6-liter plastic beaker
Chromatography column: Pharmacia Biotech XK 16/100, 26/100, or 50/100
Packing reservoir: Pharmacia Biotech RK 16/26 (for 16- and 26-mm-i.d. columns) and RK 50 (for 50-mm-i.d. column)
Chromatography pump: Pharmacia Biotech P-6000 or P-500
Injection valve (to select between sample loop and pump)
UV monitor and fraction collector
Sample loop (volume determined by size of column)

NOTE: The various components of the chromatography system (pumps, valves, monitors, and sample loops) listed separately above are supplied as components of the BioPilot chromatography system (Pharmacia Biotech), which is used to run the XK 50/100 column. The smaller XK columns (2.6 and 2.5 cm i.d.) are run using the FPLC chromatography system (also from Pharmacia Biotech), which is designed for small- to medium-scale work. For further details on this equipment see the manufacturer's literature (e.g., *Process Products*, Pharmacia Biotech).

NOTE: Perform steps 1 to 11 at room temperature. After the column is packed, equilibrate and elute at 4°C.

Pack the column

1. Wash the gel-filtration medium in a large plastic beaker with 5% ethanol. Let the medium settle and adjust the volume of liquid to give a gel slurry concentration of 65% to 75%.

   The XK 16/100, 26/100, and 50/100 columns are 100 cm long and have inner diameters of 16, 26, and 50 mm, respectively. Hence, for an XK 50/100 column, column volume = radius (2.5 cm)^2 × 3.1416 × bed height (97 cm) ≅ 1900 ml, and ~2 liters preparative-grade Superdex 200 is required. To pack this column, the gel medium is suspended in 5% ethanol to give a total volume of 3 liters which corresponds to ~70% gel slurry (it should be noted that the RK 50 reservoir has a capacity of 1 liter, so the 3 liters of gel slurry can be poured in a single operation).

2. Fix the chromatography column in an upright position, using a level to adjust the position. Attach the packing reservoir.

3. Add sufficient 5% ethanol to displace the air from a few centimeters of the bottom of the column. Clamp off the bottom of the column.

4. Gently mix the gel-filtration medium in the plastic beaker to an even slurry of 70% medium suspended in 5% ethanol.

5. Degas the suspension 5 to 10 min using a vacuum flask and laboratory vacuum.

   The ethanol is included to reduce the surface tension and density of the solvent, thus allowing air bubbles that form to rise to the surface more quickly.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

6. Carefully pour the slurry of medium into the column, introducing material along the side of the column to avoid creating air bubbles.

7. Let the column stand 5 min and then unclamp the bottom of the column.

8. Attach the chromatography pump to the packing reservoir and pump 5% ethanol (degassed) into the column at an appropriate flow rate (based on manufacturer's instructions). Pack the column at a pressure greater than the pressure at which the column will be run (up to twice as high), but not greater than the maximum pressure rating of the column.

   The XK 50/100 column (rated to 0.5 MPa) is packed at ~20 to 30 ml/hr and ~0.4 MPa.

9. After the medium has settled, turn off the pump and close the bottom of the column. Pipet fluid from the reservoir and remove the reservoir.

   Once the column has been packed, be careful to prevent air from entering the column bed. Air will disturb the bed and reduce the column separation resolution.

10. Attach the column top adapter to the column. Place the top of the adapter onto the top of the packed medium and gently compress the medium.

11. Reattach the pump to the column and wash the column with water at a flow rate that will generate the maximum pressure to be used. If the medium continues to settle, readjust the top adapter to maintain a firm fit against the gel.

From this point onward, perform all steps at 4°C.

Equilibrate the column

12. Equilibrate the column with at least 1 column volume of gel-filtration buffer.

    Although the proteins were extracted with buffer containing 8 M guanidine·HCl (see Basic Protocol 1), the gel-filtration buffer contains only 4 M guanidine·HCl. The concentration is reduced to allow faster flow rates and for reasons of economy. Most proteins remain unfolded at the lower guanidine·HCl concentration. If, however, the protein elutes in an anomalous manner (e.g., in more than one peak or at an elution position not consistent with its size), and assuming there is adequate reducing agent present, then try increasing the guanidine·HCl concentration in the gel-filtration buffer.

13. Measure the actual flow rate while running the column at a flow rate that generates a back pressure about one-half of that generated when packing the column (step 8).

    For an XK 50/100 column packed using Superdex 200 at 0.4 MPa, a running pressure of ~0.2 MPa is used, which generates flow rates of 5 to 10 ml/min that are equivalent to linear flow rates of 15.3 to 30.6 cm/hr. The linear flow rate equals the flow rate (ml/hr)/cross-sectional area (cm^2). At these flow rates it takes between 3 and 6 hr to complete the chromatography.

14. Connect tubing from the end of the column to the UV monitor and the fraction collector.

Apply the sample

15. Load the sample loop with the guanidine·HCl extract to be separated.

    Avoid loading a sample volume >5% of the total column volume; the optimum sample size is 2% (~40 ml for the XK 50/100 column). The sample consists of washed pellets extracted with guanidine·HCl (see Basic Protocol 1). A sample size of 40 to 50 ml is usually derived from ~50 g wet weight cells. With smaller sample sizes, use columns with proportionally smaller diameters (e.g., XK 16/100 or 26/100 columns). If pur-

BOSTON UNIVERSITY MEDICAL CENTER LIBRARY

chase of only one column is possible, a 2.5 × 100–cm size is a good compromise for variable sample loading.

16. Monitor column effluent with the UV monitor and collect fractions with the fraction collector.

    For an XK 50/100 column, collect 15- to 20-ml fractions in 16 × 20–mm culture tubes.

    The eluent from the column is usually monitored at 280 nm or, if the protein has a particularly low extinction coefficient, at 230 nm (guanidine·HCl strongly absorbs below 225 nm). For an XK 50/100 column, fractions need only be collected after ~500 ml of elution. The excluded volume (void volume) is ~570 ml. Run one column volume (1900 ml) to ensure all of the load material is eluted from the column.

17. Prepare the fractions to be assayed for SDS-PAGE (see Support Protocol and *UNIT 10.1*).

SUPPORT PROTOCOL

PREPARATION OF SAMPLES CONTAINING GUANIDINE HYDROCHLORIDE FOR SDS-PAGE

Because guanidine·HCl forms a precipitate with SDS, it is necessary to remove the former before carrying out SDS-PAGE. Protein in column fractions is separated from guanidine·HCl by precipitation using 90% ethanol (Pepinsky, 1991).

Materials

Sample containing the protein of interest
100% ethanol, 0° to 4°C
1× SDS sample buffer (*UNIT 10.1*)

Gilson Pipetman (Rainin Instrument)

Additional reagents and equipment for gel electrophoresis (*UNIT 10.1*)

1. Pipet 25 µl sample containing the protein of interest into a 1.5-ml microcentrifuge tube.
2. Add 225 µl cold (0° to 4°C) ethanol to the sample in the tube.

   The final ethanol concentration is 90% by volume.

3. Mix the sample and ethanol well. Chill 5 to 10 min at –20°C or colder (e.g., –80°C).
4. Microcentrifuge the sample 5 min at maximum speed (~15,000 × *g*), 4°C. Carefully withdraw the supernatant and retain the pellet.

   The pellet may be difficult to see. Be careful not to draw the pellet out of the microcentrifuge tube with the supernatant.

5. Suspend the pellet with 250 µl cold 90% (v/v) ethanol. Mix thoroughly using a vortex mixer.

   The 90% ethanol is made by mixing 225 µl ethanol and 25 µl H_2O.

6. Microcentrifuge the sample 5 min at maximum speed, 4°C. Carefully pipet off the supernatant and suspend the pellet in 25 µl of 1× SDS sample buffer.

   Some proteins are more difficult than others to suspend from an ethanol precipitate. Electrophoresis sample buffer containing 8 M urea is helpful for such proteins (UNIT 10.1). Sonication with a microtip probe can also be used to disperse the sample. A volume of sample buffer >25 µl may be required in this case (e.g., 50 µl), and great care must be taken to prevent foaming of the sample caused by excessive sonication power.

RCU IASI CENTRAL UNIVERSITY LIBRARY

7. Heat the sample 3 to 5 min at 90° to 100°C. Load on an SDS-polyacrylamide gel (*UNIT 10.1*).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Extraction buffer

50 mM Tris·Cl, pH 7.0
5 mM EDTA
8 M guanidine·HCl (764 g/liter)
5 mM DTT (770 mg/liter)

If the buffer is cloudy, filter through a 0.45- to 0.5-μm filter (the solution should be clear if high-quality guanidine·HCl—e.g., ultrapure grade, ICN Biomedicals—is used; see *APPENDIX 3A*). Buffer can be stored minus DTT at least 1 month at 4°C.

Gel-filtration buffer

50 mM Tris·Cl, pH 7.5
4 M guanidine·HCl (382 g/liter; ultrapure, ICN Biomedicals)
5 mM DTT (770 mg/liter)

Buffer can be stored minus DTT at least 1 month at 4°C. Filter (as for extraction buffer; see recipe) and degas before use.

Higher concentrations of guanidine·HCl (up to 8 M) may be required for some proteins (see comment at step 12).

Lysis buffer

100 mM Tris·Cl, pH 7.0
5 mM EDTA
5 mM DTT (770 mg/liter)
5 mM benzamidine·HCl (780 mg/liter)
Prepare immediately before use

The Tris·Cl and EDTA are diluted from concentrated stock solutions. The other components are added to the diluted buffer before use.

Wash buffer

100 mM Tris·Cl, pH 7.0
5 mM EDTA
5 mM DTT (770 mg/liter)
2 M urea (120 g/liter; ultrapure, ICN Biomedicals)
2% (w/v) Triton X-100 (20 g/liter; Calbiochem-Novabiochem)

Add DTT, urea, and Triton X-100 to the other components directly before use. Prepare this buffer in two forms: one with and one without the urea and Triton X-100 (the latter for use in Basic Protocol 1, step 9).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

The decision of whether to work with insoluble recombinant protein or to put more effort into generating soluble protein (e.g., by modifying the expression vector or changing the host strain and fermentation conditions) can be dictated by the nature of the protein. A small protein (10 to 17 kDa) with only one or two cysteine residues might be expected to fold in reasonable yield from extracted inclusion bodies. Larger proteins (>25 kDa) with many cysteine residues may be more problematical, and lower folding yields can normally be expected. In the latter case, if only small amounts of material are needed then yield is not such an important issue.

It should be emphasized that, unless proved otherwise, a protein folded from insoluble inclusion bodies can be expected to have the same structural and conformational integrity as the same protein directly purified from soluble extracts (also see *UNIT 6.1*). It is similarly true that a purified soluble protein can be denatured and renatured (reversible denaturation) without structural or conformational modifications (reviewed by Anfinson, 1973; Ghelis and Yon, 1982).

How inclusion bodies are formed

Inclusion bodies are noncrystalline, amorphous structures; however, there is some evidence that the constituent densely packed proteins may have nativelike secondary structures (Oberg et al., 1994). This suggests that the aggregates are formed by the association of partially folded protein or so-called folding intermediates. Furthermore, the aggregation appears to take place late, rather than early, in the folding pathway.

Although inclusion body formation and its prevention are of great academic and commercial interest, the pragmatic situation is clearly summarized by Seckler and Jaenicke (1992), who state: "That inclusion bodies are aggregates of otherwise intact polypeptides in nonnativelike conformations has been proven repeatedly by the successful refolding of active proteins after dissociation of the aggregates by chemical dissociation." It must be noted that the work of Oberg et al. (1994) does not invalidate this statement. It is simply a matter of terminology: the protein in aggregates may have nativelike secondary structures, but compared with soluble folded protein it must still be considered nonnative.

The propensity of a protein to form inclusion bodies is not related to the presence of sulfhydryl residues, as proteins without sulfhydryls still form such aggregates. Furthermore, where carefully studied, it has been found that aggregates derived from proteins that in their native state contain disulfides consist mainly of reduced protein (Langley et al., 1987). This also supports the view that inclusion bodies are not aggregates of completely unfolded protein because, if so, random disulfide bond formation would occur despite the reducing environment of the *E. coli* cytoplasm. Regardless of the sulfhydryl state in the aggregates, once the protein has been solubilized, reducing agents must be included to prevent the formation of nonnative disulfide bonds.

Extracting inclusion body protein

Proteins are extracted from inclusion bodies using strong protein denaturants. Protein denaturation can be induced by the following solvent conditions or reagents.

1. *pH.* Protein denaturation occurs because of the ionization of side chains. Generally, proteins retain less residual structure (are more denatured) when exposed to high pH (e.g., >10.5) compared to low pH (<4.5). Acidic or basic pH may be used in conjunction with urea or inorganic salts. An example of acidic pH extraction is given in *UNIT 6.5*.

2. *Organic solvents.* Examples of organic solvents include ethanol and propanol. Generally, proteins are not completely unfolded in these solvents. Organic solvents are infrequently used for the primary extraction of inclusion bodies but can be useful cosolvents to enhance protein folding (Chang and Swartz, 1993).

3. *Organic solutes.* Examples of organic solutes include guanidine·HCl (used at 6 to 8 M; effectiveness independent of solvent pH but temperature-dependent) and urea (used at 6 to 9 M; effectiveness modulated by pH, ionic strength, and temperature). Organic solutes are the most versatile and most commonly used denaturants for solubilizing inclusion bodies. The denaturing power of guanidinium salts increases in the order $Cl^- < Br^- < I^- < SCN^-$.

4. *Detergents.* The most common examples of protein-denaturing detergents are sodium dodecyl sulfate (SDS, an anionic detergent) and cetyltrimethylammonium bromide (CTAB, a cationic detergent). Both are effective protein denaturants, but they are not nor-

PCU IASRI CENTRAL UNIVERSITY LIBRARY

mally used to solubilize inclusion bodies as they interfere with recovering folded protein in reasonable yields (e.g., see van Kimmenade et al., 1988).

5. *Inorganic salts.* Salts at high concentration (>1 M) often denature proteins. The denaturing power increases in the following order for anions: $SO_4^{2-} < CH_3COO^- < Cl^- < Br^- < ClO_4^- < SCN^-$; for cations, the order is $(CH_3)_4N^+$, NH_4^+, K^+, $Na^+ < Li^+ < Ca^{2+} < Gdn^+$ (reviewed by von Hippel and Schleich, 1969). The organic guanidinium ion (Gdn) is included for comparative purposes. Salts have not been widely used for solubilizing inclusion bodies. On the other hand, they are frequently used for selective extraction of extrinsic membrane proteins and in principle should be useful for preextracting inclusion bodies. Denaturing salts such as KSCN or LiBr are sometimes referred to as chaotropes (tending to disorder).

6. *Temperature.* Thermally induced extraction is rarely used as it often results in irreversible protein denaturation (Zale and Klibanov, 1986). However, denaturants such as guanidine·HCl and urea are more effective at elevated temperatures (e.g., 37° to 60°C). Care must be taken when heating urea solutions because of the increased rate of cyanate formation, which will covalently modify amino groups on the protein, especially at pH >6.

A comparison of some protein denaturants and their relative effectiveness is described by Pace and Marshall (1980, and references cited therein). Monera et al. (1994) provide an explanation of why guanidine·HCl is usually two to three times per mole (2- to 3-fold) more effective than urea at unfolding proteins. The mechanism of protein denaturation is reagent-specific; this topic has been reviewed by Tanford (1968), Ghelis and Yon (1982), and Creighton (1993).

The extraction conditions can be empirically determined on a small scale (<1 ml) by taking aliquots of washed pellet suspension and microcentrifuging. If preweighed microcentrifuge tubes are used, the extraction conditions can be related to a given wet weight of pellet. A fixed volume of extraction buffer is added plus a reducing agent (5 to 20 mM DTT). Initially, guanidine·HCl (4 to 8 M) or urea (6 to 9 M) should be tried. The pellets are resuspended by vortexing, then incubated for various times at different temperatures. (Suspension of the pellet is aided by brief sonication or by using a tissue homogenizer.) After equilibration, solutions are clarified either by filtration or by centrifugation. For the latter, a benchtop ultracentrifuge such as the Beckman XL-100 or Airfuge is ideal. The comparative effectiveness of extractants can determined by measuring the amount of protein solubilized from the washed pellets using standard protein estimation methods and/or SDS-PAGE (*UNIT 10.1*).

It should be noted that solubilized protein must at some stage be folded into the native conformation; hence, the most effective solubilizing conditions might not necessarily be the best, especially if they cause irreversible denaturation. Irreversible denaturation appears to result from chemical modification of the protein and is induced by such factors as high temperature, extremes of pH, and tight binding of denaturants such as SDS.

The extraction process (Basic Protocol 1) should result in protein that is both monomeric (assumed to be unfolded) and contains cysteine residues in the reduced state. This provides a defined starting point from which to develop a reproducible folding protocol or from which to further purify the protein by, for example, gel filtration under denaturing conditions (Basic Protocol 2). Extraction of inclusion bodies with some denaturants, such as urea, may not completely convert the protein to monomers, resulting in physically heterogeneous mixtures. As pointed out in *UNIT 6.1*, extraction can be accomplished with strong denaturants (e.g., guanidine·HCl) that can then be exchanged by dialysis or gel filtration for weaker ones (e.g., urea). This particular solvent exchange often results in much better yields of folded protein compared to that obtained by the direct removal of guanidine·HCl (e.g., *UNIT 6.5*).

Gel-filtration chromatography (Basic Protocol 2) is not commonly considered a high-resolution separation technique. However, as the recombinant-derived protein content of well-prepared washed pellets will be >50% of the total (see Fig. 6.3.1, lanes h and i), only a 2-fold purification is required to obtain pure protein.

Protein extracted with guanidine·HCl in the presence of reductant will ideally be in a random coil conformation with all sulfhydryl residues in the reduced state. Under such conditions, the order in which proteins elute from a gel-filtration matrix in guanidine·HCl can be directly correlated with molecular size (Mann and Fish, 1972).

Selection of the proper chromatography resin is critical for success (for detailed discussion, see Critical Parameters and Troubleshooting). The main disadvantage of gel filtration in

B. C. U. „... SCU" IAŞI

guanidine·HCl, especially when using some of the (soft) agarose-based resins (e.g., Sepharoses and Bio-Gels), is that flow rates can be very slow due to the high viscosity of the gel-filtration buffer. Basic Protocol 2 uses a Superdex matrix that permits fast flow rates in the presence of high guanidine·HCl (or urea). A column size of 5 × 100 cm allows ~40 to 50 ml guanidine·HCl-containing extract to be processed. For smaller sample sizes, columns with proportionally smaller diameters are used.

Because the proteins separated by gel filtration in the presence of guanidine·HCl are unfolded, they have no biological activity, so the column fractions are usually assayed by SDS-PAGE. Direct addition of SDS to samples containing guanidine·HCl results in the formation of the insoluble guanidinium salt of dodecyl sulfate. The guanidine·HCl must therefore be removed before addition of the sample buffer used for SDS-PAGE analysis (*UNIT 10.1*). The approach used in Support Protocol 1 takes advantage of the fact that guanidine·HCl is soluble in 90% ethanol whereas the protein is not. An alternative is to simply remove the guanidine·HCl by dialysis. Small samples (10 to 100 µl) can be dialyzed using one of various microdialyzing systems commercially available (e.g., the Microdialyzer System 100 from Pierce, which allows simultaneous dialysis of up to 12 samples) or a simple homemade device that uses modified microcentrifuge tubes (Falson, 1992).

Critical Parameters and Troubleshooting

Breaking cells and preextracting inclusion bodies

Once it has been established that the recombinant protein is insoluble (*UNIT 6.1*), extra care should be taken to ensure complete cell lysis. If unbroken cells are present in the low-speed pellets, they will leach contaminants when the pellets are extracted with strong protein denaturants. Proper operation of the French press is described in *UNIT 6.2*. Other major sources of contamination are outer membrane and cell wall material, most of which can be preextracted with detergents such as Triton X-100 (0.5% to 5%) and urea (1 to 4 M).

The proper concentration of urea in the wash buffer is determined empirically in small-scale extractions. It is probably safe to use 2 M urea for most aggregates. The effectiveness of the wash process at removing the cell wall and outer membrane material can be improved by increasing the volume of wash buffer (see Basic Protocol 1, step 6: use 10 ml instead of 6 ml buffer per gram cells). Treatment of the cells with lysozyme prior to homogenization with the French press will also help remove this material (*UNIT 6.5*, Basic Protocols 1 and 2).

Although insoluble proteins are less susceptible than native proteins to proteolysis, inclusion of EDTA and a serine protease inhibitor such as benzamidine·HCl or AEBSF is recommended. The choice of cell lysis buffer is not critical; however, the buffer pH should be >6.5, as many soluble *E. coli* proteins precipitate at slightly acidic pH values.

After preextraction, wash the aggregates with buffer alone to remove excess detergent. When attempting to develop a reproducible folding protocol, note that varying amounts of detergent carryover may influence the outcome. In fact, nonionic detergents can be used as cosolvents to aid folding (*UNIT 6.1*); even so, it is important to know how much detergent is present.

Extracting inclusion bodies with guanidine·HCl

The washed pellets containing the insoluble inclusion body proteins are extracted with guanidine·HCl (see Basic Protocol 1, step 10). To obtain complete dissolution of the pellets, some form of mechanical dispersion is often required. Homogenization at room temperature with a tissue grinder (as described) is often adequate; however, sonication should be used if the pellet is especially recalcitrant. Heating the solution will also aid protein solubilization; 10 to 15 min at 50° to 60°C is usually a good starting point. Excess heating, whether direct or caused by sonication without adequate cooling, should be avoided. It is worth remembering that if extraction involves extremes of heat and or pH, such conditions will favor deamidation of Asn and may also promote cleavage at the labile Asp-Pro bond and other chemical modifications of the protein (Zale and Klibanov, 1986).

It is best to use the extract directly after preparation rather than storing it frozen. In general, an unfolded protein in solution is much more susceptible to proteolytic degradation and chemical modification than its native (folded) counterpart.

Selecting medium and column for gel filtration in guanidine·HCl

Selecting the proper gel-filtration resin is one of the most critical steps. First of all, the

PCU IAS/CENTRAL UNIVERSITY LIBRARY

Table 6.3.1 Gel-Filtration Matrices Suitable for Use with Solutions Containing Guanidine Hydrochloride

Matrix[a]	Mass range (kDa)		Reference
	Native proteins	Unfolded proteins[b]	
Sepharose CL-6B	10-4,000	1-80	Mann and Fish (1972)
Bio-Gel A-5m	10-5,000	1-80	Mann and Fish (1972)
Sepharose CL-4B	60-20,000	10-300	Mann and Fish (1972)
Sephacryl S-100 HR	1-100	<1-30[c]	—
Sephacryl S-200 HR	5-250	1-50	Belew et al. (1978)
Sephacryl S-300 HR	10-1,500	1-100[c]	—
Sephacryl S-400 HR	20-8,000	1->100[c]	—
Superdex 75	3-70	<1-25	I.P. and P.T.W. (unpub. observ.)
Superdex 200	10-600	1-80	I.P. and P.T.W. (unpub. observ.)

[a]All resins are from Pharmacia Biotech except Bio-Gel A-5m, which is from Bio-Rad. The Sepharose and Bio-Gel matrices are normally run under low pressure; all other resins can be run under low or medium pressure. Medium pressure is achieved using one of the chromatography pumps indicated in Basic Protocol 2; the pumps are normally included in the Pharmacia Biotech FPLC or BioPilot systems.

[b]Data on the fractionation range in the unfolded state refer to proteins unfolded with guanidine·HCl; however, the guidelines also apply to proteins unfolded and eluted with urea (assuming they are random coils).

[c]Estimates based on fractionation range for native proteins.

chromatography medium must have high chemical stability that permits prolonged exposure to high concentrations of denaturing solvents such as guanidine·HCl and urea. It must also have high physical strength that allows high flow rates with viscous solvents. Suitable media are listed in Table 6.3.1. The Sepharoses and Bio-Gels are cross-linked agaroses, Sephacryl is dextran cross-linked with bisacrylamide, and Superdex is cross-linked agarose with bonded dextran. All the matrices listed are stable in high concentrations of urea and guanidine·HCl. Details on the properties of various media can be obtained from the manufacturers.

Second, selecting a medium with the proper pore size is important for good separation. Manufacturers provide information about the separation ranges of their products. These ranges are determined for proteins in an aqueous buffer. In guanidine·HCl, denatured proteins (assuming they are random coils) will have a radius 3- to 4-fold greater than an average folded globular protein of the same mass. Thus, the mass range values for unfolded proteins in Table 6.3.1 should be used rather than the native protein values.

Third, the pressure rating of the column must be adequate. Because 4 M guanidine·HCl has a higher viscosity than most aqueous buffers, higher pressures are needed to achieve the same flow rates. The column should be capable of safely withstanding moderate pressures of around 0.5 MPa (5 bar) or ~75 lb/in^2 (the XK columns are rated to 0.5 MPa). The proteins separate as they pass through the column, so column length is important. A column 60 to 100 cm in length will generally be suitable. Some manufacturers offer prepacked columns with a choice of media. If column packing is to be done in the laboratory, then selecting a column with a packing reservoir is helpful.

Prepacked columns (diameters 1.6, 2.6, 3.5, and 6.0 cm and length 60 cm) prepacked with Sephacryl resin (Sephacryl S-100, S-200, or S-300) or Superdex resin (Superdex 200 and 75) are commercially available from Pharmacia Biotech. Superdex 200 and 75 PG (preparative grade) 60/600 columns (6.0 cm i.d. × 60 cm length) are used with the BioPilot system (Pharmacia Biotech) either individually or sequentially to obtain, in many cases, pure protein (for example, see Fig. 6.3.1).

Performing gel filtration in guanidine·HCl

The general guidelines and critical parameters that apply to gel-filtration chromatography in general and are detailed elsewhere (*UNIT 8.3*) also apply here. The main distinction from typical gel filtration is the use of a highly

BOURNEMOUTH UNIVERSITY LIBRARY

viscous solvent, which requires robust matrices to obtain reasonable flow rates. On the other hand, it should be noted that the presence of high concentrations of guanidine·HCl gives higher-resolution separations than those obtained with normal solvents because the high viscosity of the solvent reduces diffusion, which otherwise tends to broaden elution peaks.

It is important to ensure that all connections and tubing are leak-free. Concentrated guanidine·HCl is highly corrosive to metal and electronic components.

The sample derived from extraction of highly aggregated protein must be clarified before being applied to the column. Furthermore, to prevent the sample from precipitating near or at the top of the column, it should be equilibrated to the same temperature as the column matrix and buffer. This is especially important if chromatography is carried out at 4°C and the sample was heated during extraction.

The column must be equilibrated with freshly made buffer before use. The column buffer contains the sulfhydryl reagent DTT, which will oxidize over time, thus lessening its effectiveness and increasing the background UV absorbance of the buffer.

If the recombinant protein has been effectively solubilized, it will be physically homogeneous and its elution position will be proportional to its mass. If the protein elutes anomalously (i.e., is located in more than one peak), there are several possible reasons.

1. *Incomplete protein solubilization.* Try more drastic extraction conditions—in particular, increase the denaturant concentration and temperature during extraction. Use a stronger chemical denaturant, especially if urea was originally used.

2. *Aggregation/precipitation of protein(s) in the column.* Ensure that the concentration of guanidine·HCl in the column buffer will maintain the solubility of all proteins during chromatography. The column can be run at room temperature (as opposed to 4°C), and reasonable flow rates can still be obtained with buffers containing 6 to 7 M guanidine·HCl (4 M is used in Basic Protocol 2).

3. *Conformational heterogeneity.* Related to the above two situations is the case where the protein, although monomeric, contains both unfolded and partially folded species. The folded protein will elute later (appear smaller) than the unfolded protein. Conformational heterogeneity is usually not a major problem in high (>4 M) guanidine·HCl concentrations.

4. *Oxidation of sulfhydryl residues.* The formation of intermolecular disulfide bonds is indicated by the appearance of protein eluting earlier than expected and corresponds to the formation of multimers (dimers, trimers, etc.). Disulfide interchange occurs more readily at high pH; thus, use of a slightly acidic buffer (pH 6.0 to 5.0) and inclusion of fresh reductant in both the sample and column buffer are recommended. For analytical separations, cysteine residues are normally capped by alkylation with iodoacetamide; cysteines can also be reversibility modified by sulfitolysis (Glazer et al., 1975).

5. *Proteolysis.* (a) The sample contains partially proteolyzed protein, with proteolysis by *E. coli* proteases occurring before or after cell breakage: In fully denatured protein, it should be possible to separate clipped protein from unmodified protein if there is sufficient mass difference. Once the protein has been separated by gel filtration, it should be checked by mass spectrometry to confirm its integrity. If the protein is not fully denatured, clipped protein may have similar elution properties as unmodified protein; however, the processing will be revealed when the protein is boiled in SDS and analyzed by gel electrophoresis.

(b) Where the recombinant protein is itself a protease: As pointed out by Fish et al. (1969), "Since denaturation is a kinetic process, the dissolution of proteases in guanidine·HCl might result in a condition where the fraction of the protein not denatured at a given instance may digest the unfolded protein leading to low molecular weight estimate." The converse of this situation is also true: if proteases are extracted under conditions where they are fully denatured and, thus, inactive (e.g., 8 M guanidine·HCl), changes in the solvent conditions (e.g., reducing the guanidine·HCl concentration or exchanging for urea) might lead to a fraction of the protein folding into an active conformation. If autolytic processing is suspected, include an appropriate inhibitor or choose a solvent pH where the enzyme has minimal activity.

Ancipitated Results

Cell lysis and preparation of washed pellets

Pelleted aggregates after washing contain ~30% dry weight, of which 90% is protein. SDS-PAGE of a typical washed pellet preparation (Fig. 6.3.1, lanes h and i) indicates that recombinant bovine growth hormone (21 kDa) makes up >60% of the total protein. The

RCU IASI CENTRAL UNIVERSITY LIBRARY

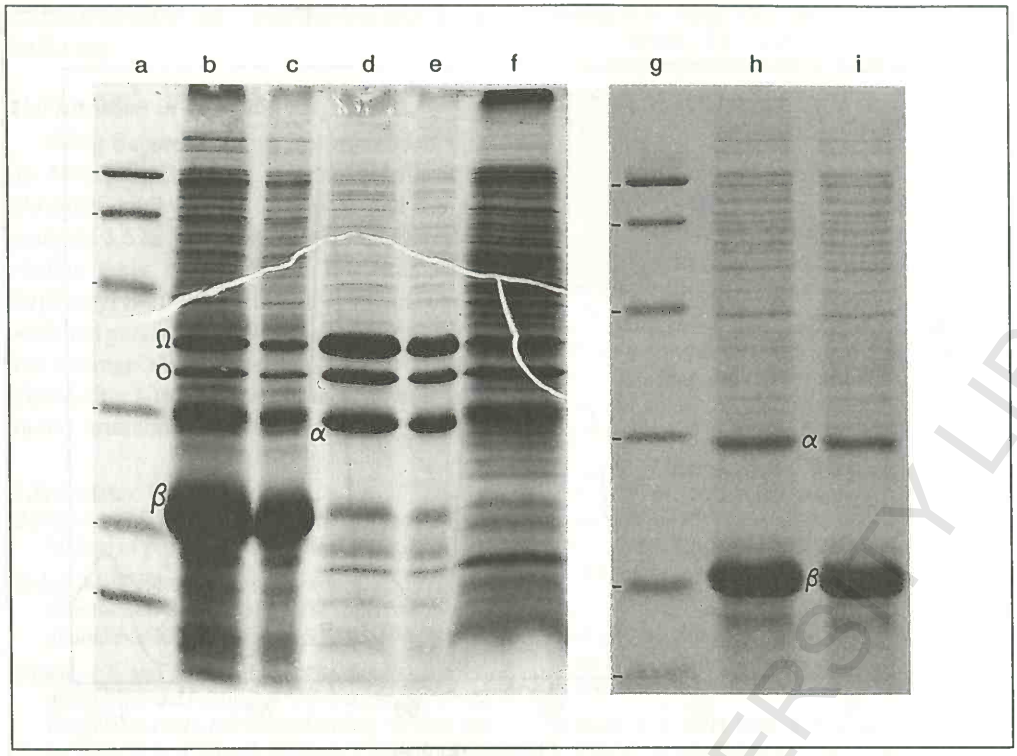


Figure 6.3.1 Analysis by SDS-PAGE of fractions from low-speed centrifugation of *E. coli* cell lysates containing aggregated bovine growth hormone. A 12.5% acrylamide gel of dimensions 12 cm × 16 cm × 1.5 mm was used with the Laemmli buffer system (*UNIT 10.1*). Lanes a and g contain standard proteins (low-range standards, Bio-Rad) in order of increasing migration distance: phosphorylase *b* (97.4 kDa), bovine serum albumin (66.2 kDa), hen egg white ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), and hen egg white lysozyme (14.4 kDa). After low-speed centrifugation of the clarified lysate and of the washed pellet homogenate (see Basic Protocol 1, steps 5 and 7), the supernatants will be cloudy (lane f) and the pellets usually consist of two layers (see Fig. 6.1.5). The bottom layer is inclusion body protein plus unbroken cells (lanes b and c) and the top layer consists of outer membrane and peptidoglycan fragments (lanes d and e). The outer membrane proteins OmpA (35 kDa) and OmpF/C (38 kDa) are indicated by ω and o, respectively. After the washing steps, the growth hormone (marked β, 21 kDa) is the major constituent (lanes h and i) together, in this example, with another plasmid-encoded protein, namely kanamycin phosphotransferase (marked α, 30.8 kDa), the product of the gene conferring resistance to the antibiotic kanamycin.

washed pellets analyzed in Figure 6.3.1 are typical starting materials for the protein folding and purification described in *UNIT 6.5* (Basic Protocol 1). The multiple bands in the background are either derived from unbroken cells (most likely) or are *E. coli* cytoplasmic proteins coprecipitated or trapped during aggregate formation. The gel analysis indicates the presence of another plasmid-encoded protein, namely, kanamycin phosphotransferase (30.8 kDa), a product of the gene conferring resistance to the antibiotic kanamycin (Kane and Hartley, 1991). The washing procedure described only partially extracts the phosphotransferase but removes most of the outer membrane proteins. For further details, see the legend to Figure 6.3.1.

Gel filtration in guanidine·HCl

Fractionation of human immunodeficiency virus type 1 (HIV-1) protease illustrates expected results from gel-filtration chromatography. The protein is expressed in *E. coli* as a 17-kDa precursor that at some stage undergoes autolytic processing to form the mature-sized protease (10.7 kDa). The protease and undigested precursor accumulate as insoluble inclusions in the *E. coli* cytoplasm. Washed pellets prepared and extracted with 6 M guanidine·HCl as described (see Basic Protocol 1) were applied to a Superdex 200 column. The protease eluted in a single peak (Fig. 6.3.2, fractions 66 to 72), well separated from unprocessed protein (e.g., fraction 60) and larger molecular mass proteins (e.g., fractions 50 to

BOSTON UNIVERSITY LIBRARY

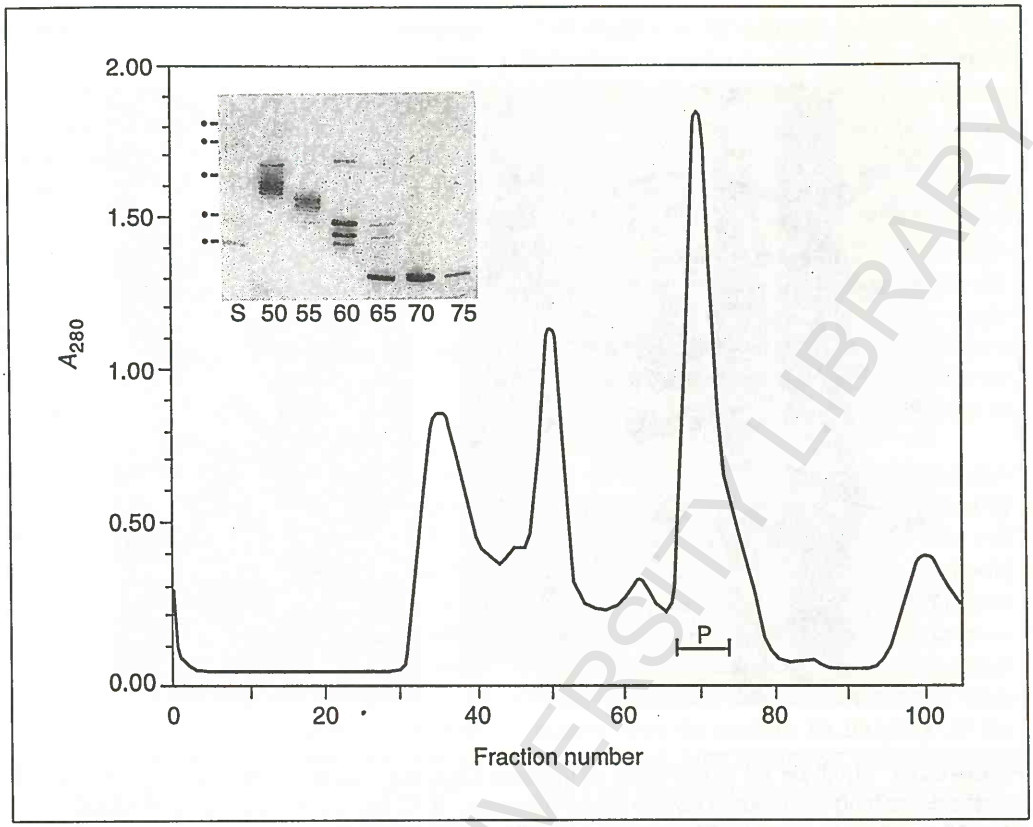


Figure 6.3.2 Gel filtration using Superdex 200 in 4 M guanidine·HCl. Column dimensions, 6 × 60 cm; buffer, 50 mM Tris·Cl (pH 7.5)/4 mM guanidine·Cl/2 mM DTT; flow rate, 5 ml/min (300 ml/hr). The sample was an extract containing HIV-1 protease, which has a mass of 10 kDa. Protein fractions 66 to 72 (pool P) was further purified under the same conditions using a Superdex 75 matrix. The inset shows SDS-PAGE analysis of selected fractions. The protein markers (lane S) correspond to standards with mass values of 66.2, 45, 30, 21.5, and 14.4 kDa, respectively (migration order top to bottom).

55). The fractions indicated in Figure 6.3.2 by P were pooled and the protein further purified by repeat chromatography on a Superdex 75 column under the same conditions. The protein at this stage is >95% pure and after solvent exchange is folded into active protein. (As it happens, the native protease is a 20-kDA homodimer that is very susceptible to autolytic digestion; hence, purification in the denatured, inactive state is highly advantageous).

In the inset of Figure 6.3.2, analysis by SDS-PAGE of fraction 60 illustrates incomplete dissociation of the sample, a problem commonly observed when electrophoresing samples precipitated with ethanol and trichloroacetic acid (TCA). If necessary, repeat the analysis using SDS sample buffer (*UNIT 10.1*; see Support Protocol).

The fractionation of HIV-1 protease is somewhat of a best-case scenario due to the small size of the protein. Usually it is not possible to purify the recombinant protein completely, especially for proteins close in size to the main *E. coli* contaminants (30 to 45 kDa). In general, small proteins (<25 kDa) are best suited for the method; however, any protein will be purified to some extent, and further purification can be attempted using some of the methods discussed in *UNIT 6.1*. Moreover, partial purification may be all that is required to enhance protein folding significantly.

Recovery of recombinant protein from gel-filtration chromatography is close to quantitative, and the purified (or partially purified) protein can be folded or can be stored in aliquots at –80°C. On thawing the protein, add fresh DTT (an additional 1 or 2 mM) and warm the solution at 37° to 40°C for several minutes before use. The protein concentration can be conveniently estimated by UV spectroscopy.

Time Considerations

Cell lysis and preparation of washed pellets

French press lysis of ~150 to 200 ml cell suspension will take ~30 to 35 min. Preparation

RCU IAS CENTRAL UNIVERSITY LIBRARY

and extraction of the washed pellet take about half a day.

Gel filtration in guanidine·HCl

Using Superdex resins in conjunction with an FPLC or BioPilot system, gel-filtration chromatography takes 3 to 6 hr and the gel analysis 1.5 to 2 hr. With low-pressure resins, elution takes much longer. For example, a Sephacryl S-300 column (2.5 × 98 cm) eluted with 6 M guanidine·HCl runs at ~25 to 30 ml/hr. An average-sized protein (15 to 30 kDa) is eluted after 10 to 12 hr. A column can be loaded in late afternoon and run overnight.

Literature Cited

Anfinson, C.B. 1973. Principles that govern the folding of protein chains. *Science* 181:223-230.

Belew, M., Fohlman, J., and Janson, J.-C. 1978. Gel filtration on Sephacryl S-200 superfine in 6 M guanidine·HCl. *FEBS Lett.* 91:302-304.

Chang, J.Y. and Swartz, J.R. 1993. Single-step solubilization and folding of IGF-1 aggregates from *Escherichia coli*. *In* Protein Folding: In Vivo and in Vitro (J.L. Cleland, ed.) pp. 178-188 (ACS Symposium Series No. 526). American Chemical Society, Washington, D.C.

Creighton, T.E. 1993. Proteins: Structures and Molecular Properties, 2nd ed., pp. 293-296. W.H. Freeman, New York.

Falson, P. 1992. An efficient procedure to dialyze volumes in the range of 10-200 µl. *BioTechniques* 13:20.

Fish, W.W., Mann, K.G., and Tanford, C. 1969. The estimation of polypeptide chain molecular weights by gel filtration in 6 M guanidine hydrochloride. *J. Biol. Chem.* 244:4989-4994.

Ghelis, C. and Yon, Y. 1982. Simulation of protein folding: Studies of in-vitro denaturation-renaturation. *In* Protein Folding, pp. 225-243. Academic Press, San Diego.

Glazer, A.N., Delange, R.J., and Sigman, D.S. 1975. Modifications of sulfhydryl and disulfide groups. *In* Chemical Modification of Proteins, pp. 101-120 (Laboratory Techniques in Biochemistry and Molecular Biology Series). North-Holland Publishing Company, New York.

Kane, J.K. and Hartley, D.L. 1991. Properties of recombinant protein-containing inclusion bodies in *E. coli*. *In* Purification and Analysis of Recombinant Proteins (R. Seetharam and S.K. Sharma, eds.) pp. 121-145. Marcel Dekker, New York.

Langley, K.E., Berg, T.F., Strickland, T.W., Fenton, D.M., Boone, T.C., and Wypych, J. 1987. Recombinant-DNA-derived growth hormone from *Escherichia coli*: Demonstration that the hormone is expressed in the reduced form, and isolation of the hormone in the oxidized, native form. *Eur. J. Biochem.* 163:313-321.

Mann, K.G. and Fish, W.W. 1972. Protein polypeptide chain molecular weights by gel chromatography in guanidinium chloride. *Methods Enzymol.* 26:28-42.

Monera, O.D., Kay, C.M., and Hodges, R.S. 1994. Protein denaturation with guanidine hydrochloride provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Sci.* 3:1984-1991.

Oberg, K., Chrunyk, B.A., Wetzel, R., and Fink, A.L. 1994. Nativelike secondary structure in interleukin-1β inclusion bodies by attenuated total reflectance FTIR. *Biochemistry* 33:2628-2634.

Pace, C.N. and Marshall, H.F. 1980. A comparison of protein denaturants for β-lactoglobulin and ribonuclease. *Arch. Biochem. Biophys.* 199:270-276.

Pepinsky, B.R. 1991. Selective precipitation of proteins from guanidine hydrochloride-containing solutions with ethanol. *Anal. Biochem.* 195:177-181.

Seckler, R. and Jaenicke, R. 1992. Protein folding and protein refolding. *FASEB J.* 6:2545-2552.

Tanford, C. 1968. Protein denaturation. *Adv. Prot. Chem.* 23:122-275.

van Kimmenade, A., Bond, M.W., Schumacher, J.H., Laquoi, C., and Kastelein, R.A., 1988, Expression, renaturation and purification of recombinant human interleukin 4. *Eur. J. Biochem.* 173:109-114.

von Hippel, P.H. and Schleich, T. 1969. Ion effects on the solution structure of biological macromolecules. *Acc. Chem. Res.* 2:257-265.

Zale, S.E. and Klibanov, M. 1986. Why does ribonuclease irreversibly inactivate at high temperature? *J. Biol. Chem.* 25:5432-5444.

Contributed by Ira Palmer and
Paul T. Wingfield
National Institutes of Health
Bethesda, Maryland

BCU IAS CENTRAL UNIVERSITY LIBRARY

Overview of Protein Folding

HOW PROTEINS FOLD

Where proteins are expressed as insoluble aggregates—i.e., bacterial inclusion bodies—the obligatory first step in the recovery of functional protein is to disperse the inclusion bodies in a denaturant such as 6 M guanidine·HCl (guanidinium chloride; GdmCl), which will yield a solution of unfolded and essentially noninteracting polypeptide chains. The next problem is the one that faced Anfinsen in his pioneering work on the mechanism of protein folding—i.e., how to recover the protein in its unique, functional, three-dimensional conformation. The main factors determining the folding of proteins, and hence their three-dimensional conformation, are summarized in the following three sections.

All the Information is in the Sequence

Anfinsen (1973) showed that proteins fold up solely as a result of thermodynamic drive, which leads to the creation of the specific noncovalent interactions and disulfide bonds characteristic of the functional state. In other words, a protein folds because, under folding conditions, its native state is more stable than the unfolded state–the information determining the relative stabilities of the two states resides solely in the primary sequence. No additional energy or information, for example, in the form of a template, is required.

Other factors that may increase the yield and rate of folding such as chaperone proteins, protein disulfide isomerase (PDI), and peptidyl prolyl *cis-trans* isomerase (PPI), all depend on this thermodynamic drive for the intrinsic ability of the intact polypeptide chain to fold.

Proteins Fold Along a Limited Number of Pathways

Like a large river, proteins in the process of folding start from a large number of sources or conformations—i.e., the unfolded state(s). These rapidly condense, reducing the number of paths (intermediate states) available to the flow, until the ocean (the functional, folded state) is finally reached through something like a single pathway. As the protein folds, the number of possible intermediate states becomes increasingly limited, and further folding is thus confined to pathways by which the vast majority of conformational states potentially accessible to the polypeptide chain are precluded. In this way, proteins can fold in vivo at a rate that can keep pace with the rate of synthesis; in vitro, protein molecules can fold with half-times varying from tens to thousands of milliseconds.

Consider first proteins that do not contain disulfide bonds in the native state. Very early in the folding pathway, in less than a few milliseconds, the chain collapses to intermediates in which substantial proportions of the secondary structure and hydrophobic core are assembled. The stabilization of secondary structure elements and the packing of side chains continues until a compact intermediate termed the "molten globule" is formed. This contains essentially all the native secondary structure and, in many proteins, has been shown to accumulate to a significant extent (Ptitsyn et al., 1990). Although the topology of the molten globule is similar to that of the protein in the native state, the specific and persistent tertiary interactions characteristic of the native state are still lacking. This accounts for its ability to bind the hydrophobic probe 8-anilino-1-naphthalene sulfonic acid (ANS) and presumably also for its observed tendency to aggregate. The molten globule finally folds to the native state in a reaction that usually constitutes the rate-limiting step of folding.

Although several proteins have been shown to exhibit multiple kinetic phases and hence small energy barriers during folding, the early stages are fast ($t_{1/2}$ = 1 to 25 msec). Many of the slower steps observed in folding are rate-limited by the slow *cis-trans* isomerization of peptide bonds joining X-Pro residues (Nall, 1994). This isomerization, with half-times on the order of a minute, occurs at later stages of folding, and therefore within rather collapsed forms of the protein. The variable accessibility of the bonds involved in *cis-trans* isomerization, resulting from the collapsed state of the polypeptide chain, explains the somewhat variable success of peptidyl prolyl *cis-trans* isomerase (PPI) in catalyzing the reaction in vitro.

In comparison to these generally fast folding processes, disulfide bond–containing proteins fold somewhat more slowly in vitro. Where the noncovalent interactions are relatively strong and constitute the main factors driving the folding (at least up to a molten globule state), the stages of folding prior to disulfide formation will be similar to those described above. In some disulfide-containing proteins, however, the noncovalent interactions are weaker;

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY LIBRARY

thus the drive to shift the equilibrium toward the compact folded state must come from the formation of the native disulfide bonds. In such cases the intermediate that accumulates during folding will be relatively expanded.

Protein disulfide isomerase (PDI) is located in the lumen of the endoplasmic reticulum and is important in catalyzing formation of disulfide bridges in proteins as they fold in vivo. The isolated enzyme can, in many cases, be used to increase the rate of folding in vitro. This application of PDI has had variable success, which is accounted for (as in the case of PPI, discussed above) by the variable accessibility of the protein thiols. As PDI is a low-efficiency enzyme, large amounts are required to achieve catalysis.

Aggregation is the Main Competitor to Folding

Most proteins tend to exhibit some tendency toward association, reflecting the high proportion (usually ~50%) of their surface that is nonpolar. In the case of folding intermediates, which have somewhat higher proportions of nonpolar groups exposed to the solvent, it is not surprising that aggregation is a frequently experienced hazard. Aggregation during folding in vivo has been well-documented in the cases of phage tail spike protein mutants (Haase-Pettingell and King, 1988) and α-1-antitrypsin (Lomas et al., 1993). It has also been established that a major role of the families of chaperone proteins is to prevent aggregation from competing with folding in the cell (Hlodan and Hartl, 1994). In the case of multimeric proteins, whose subunit chains have evolved to interact strongly with each other—frequently through large nonpolar surfaces—it is not surprising to find that the occurrence of aggregation during folding is a much greater problem.

The dependence of yield on aggregation is illustrated in Figure 6.4.1.

HOW TO FOLD PROTEINS

The three principles outlined in the discussion of How Proteins Fold lead to three sets of conditions necessary for the successful recovery of functional proteins after denaturation: (1) the folded protein should be the most stable state (see All the Information is in the Sequence, above), (2) kinetic barriers that slow down or block the folding pathway should be minimized (see Proteins Fold Along a Limited Number of Pathways), and (3) intermolecular aggregation between folding intermediates should be reduced to acceptable limits (see Aggregation is the Main Competitor to Folding). These three goals will be considered as they bear on the practical problems involved in folding recombinant proteins. They will thus provide a rationale for the protocols in *UNIT 6.5*.

Stabilizing the Native Functional State

The "simple" case

Some proteins—e.g., phosphoglycerate kinase—are composed of a single intact polypeptide chain, constituting the simple case in which all the information required to define the folding and stability of the three-dimensional conformation is present in the sequence. In such a case, the native state will represent the most readily accessible stable state. Where disulfide

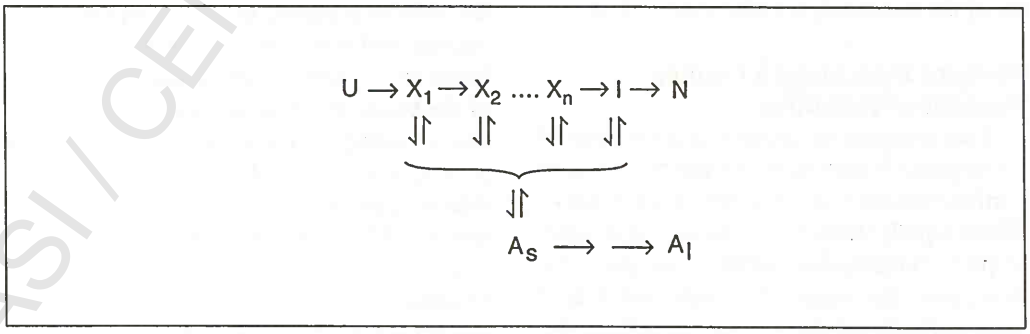


Figure 6.4.1 Illustration of the dependence of yield on aggregation. The unfolded protein (U) folds through the intermediate forms (X_1 to X_n), then through the molten globule (I) to the native state (N). Some or all of the intermediates will be capable of associating, reversibly at first, to form small aggregates (A_s). Subsequently, however, these will associate irreversibly to form larger insoluble aggregates (A_l). The rate of aggregation (r_a) will depend on the concentrations of the intermediates (and hence on the rate constants for the individual steps in the folding pathway) and also on the rate constant ($k_{i,aggr}$) for aggregation of each intermediate (X_i) according to the equation $r_a = k_{i,aggr}[X_i]$. Aggregation will therefore depend on the overall protein concentration as well as the solubility properties of each intermediate.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

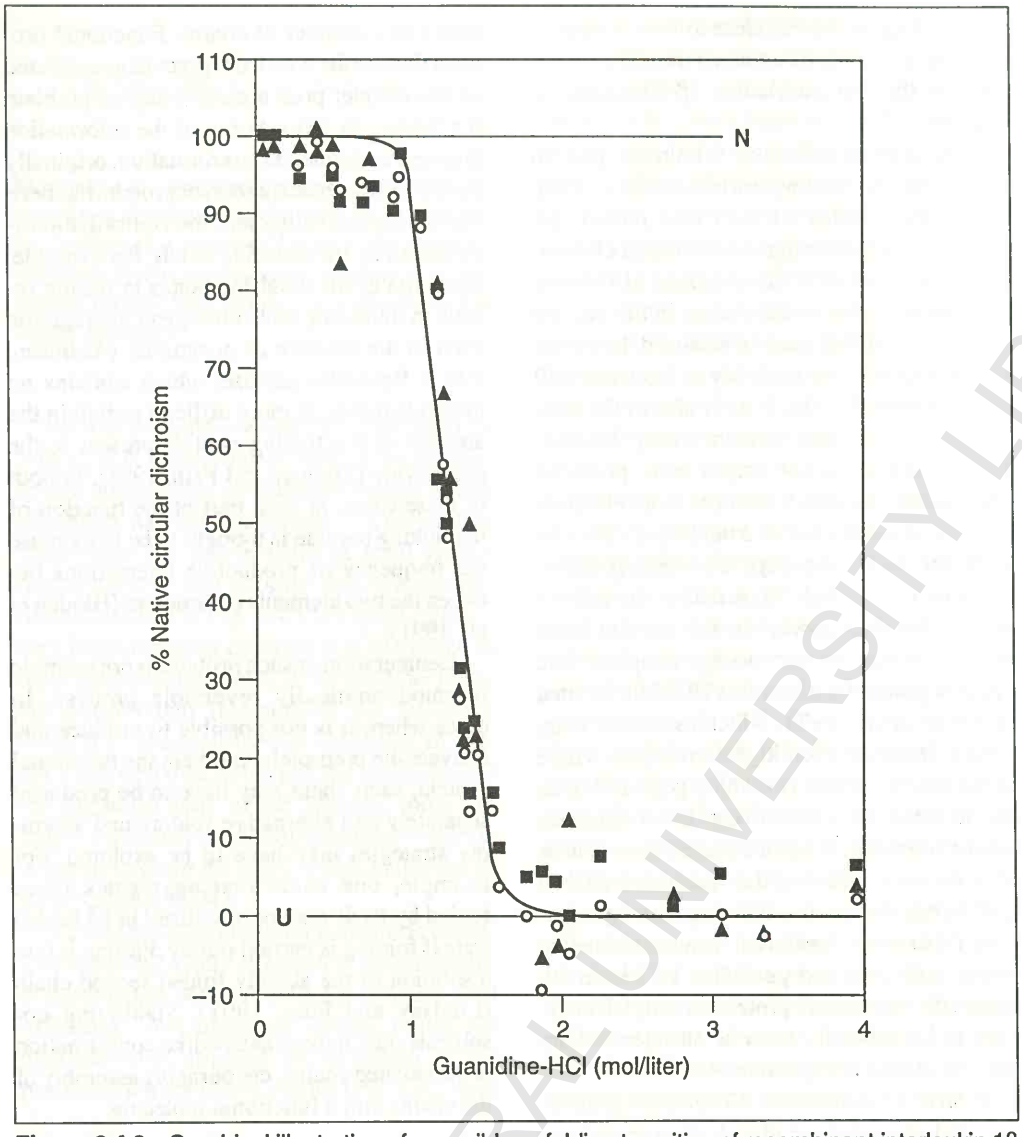


Figure 6.4.2 Graphical illustration of reversible unfolding transition of recombinant interleukin 1β in guanidine·HCl, monitored by circular dichroism (CD). Near-ultraviolet CD, which indicates changes in tertiary structure, was measured at 260 nm (triangles), 287 nm (circles), and 295 nm (squares). The plateau regions are marked for the native (top; N) and unfolded (bottom; U) states respectively, and represent regions of denaturant concentration within which there is no change in protein conformation detectable by CD. CD measurements in the far-ultraviolet range, which reflect changes in secondary structure, resulted in a curve that was coincident with this one, indicating that at equilibrium, this transition is largely two-state—i.e., only the native and unfolded states are significantly populated (Craig et al., 1987).

bonds are involved, they supplement the noncovalent interactions in contributing to the stability and specificity of the native three-dimensional conformation and the folding intermediates (see discussion of Disulfide-bonded proteins, below). Covalently bonded ligands (e.g., heme in cytochrome *c*) may also contribute to the stability of both the native and intermediate forms and therefore play an important role in folding (Roder and Elöve, 1994).

When such proteins are subjected to denaturing agent, such as guanidine·HCl, a reversible unfolding transition is normally observed (Fig. 6.4.2). For the majority of proteins, the equilibrium transition region reflects the varying proportions of the native and unfolded states only, because the concentrations of folding intermediates are rather low at equilibrium.

To fold a protein from inclusion bodies solubilized with 6 M guanidine·HCl, for example, the concentration of denaturant must be reduced to a value within the native plateau

BYU / HARVARD UNIVERSITY LIBRARY

region. It is useful therefore to have access to an experimentally determined transition curve such as that for interleukin 1β illustrated in Figure 6.4.2.

Effects of denaturants. Within the plateau region of the folding transition curve (Fig. 6.4.2), the stability of the native protein decreases with increasing concentration of denaturant. The possible disadvantage of limiting the concentration of denaturant in the interest of protein stability may be balanced, however, by the fact that the tendency to associate will be decreased. This may apply also to the transient intermediates formed during folding. Guanidine·HCl is not only a more powerful denaturant than urea (~2.5-fold more effective on a molar basis), but its solubilizing effect on different chemical groups in proteins is different from that of urea. In particular, the ratio of the solubilizing power for the peptide bond relative to that for the average nonpolar side chain is greater for guanidine·HCl than for urea (Lapanje et al., 1978; Mitchinson and Pain, 1985). Thus, for a folding intermediate, where secondary structure (in which peptide bonds are shielded from the solvent) has formed but where nonpolar groups are more exposed, urea may be more effective than guanidine·HCl in stabilizing the intermediate against aggregation (assuming denaturant concentrations at which both urea and guanidine-HCl have the same effect on overall protein stability). Hence, urea is occasionally used as an intermediate solvent during renaturation.

Effects of cosolvents. Cosolvents such as glycerol and ethylene glycol, glucose and sucrose, certain anions such as phosphate and sulfate, and certain cations such as 2-(*N*-morpholino)ethanesulfonic acid (MES) can stabilize proteins markedly against denaturation by urea or guanidine·HCl. They exhibit little or no effect on the folding rate constant, but dramatically decrease the unfolding rate constant. In certain favorable cases, the unfolded protein can be refolded simply by adding sulfate to the solution containing denaturant (Mitchinson and Pain, 1985). However these cosolvents, which act by stabilizing hydrophobic interactions, also stabilize aggregates; hence their usefulness is likely to be limited, probably to proteins with a relatively high proportion of polar residues.

Proteolytically processed proteins

Other proteins—e.g., chymotrypsin—are derived from precursors that have been processed by proteolytic cleavage and are composed of a number of chains. Functional proteins that are derived from precursors activated in this manner pose a more complex problem for folding in vitro. Some of the information required to define the conformation, originally resident in the intact precursor protein, has been lost during activation and the conformation is stabilized by the disulfide bonds. For example, reduction of the disulfide bonds in insulin results in unfolding and subsequent aggregation even in the absence of denaturant (Anfinsen, 1967). Penicillin acylase, which contains no disulfide bonds, is more difficult to fold in the absence of the linking peptide present in the proenzyme (Lindsay and Pain, 1991). In both of these cases, at least part of the function of the linking peptide is thought to be to increase the frequency of productive interactions between the two elements of structure (Hlodan et al., 1991).

Denaturation of such proteins is not a simple thermodynamically reversible process. In cases where it is not possible to produce and activate the preprotein to obtain the functional protein, each chain may have to be produced separately and alternative folding and assembly strategies may have to be explored. For example, one chain that aggregates when folded by itself may be stabilized in its folded state if folding is carried out by diluting it into a solution of the already folded second chain (Lindsay and Pain, 1991). Stabilizing cosolvents may induce native-like conformation in the isolated chains, encouraging assembly of the chains into a functional molecule.

Disulfide-bonded proteins

Disulfide bonds contribute to the stability of the native conformation to an extent that varies from one protein to another. The overall stability may be expressed thermodynamically in terms of the equilibria for the sidechain and backbone noncovalent interactions (Fig. 6.4.3) and the covalent interaction (Fig. 6.4.4) in which two thiol groups produce a disulfide bond.

$$U(-SH)_{2n} \rightleftharpoons N'(-SH)_{2n}$$

Figure 6.4.3 Equilibrium for noncovalent interactions during protein folding, where *n* is the number of disulfide bonds in the native, functional protein.

BCU IAS KENTERAN UNIVERSITY LIBRARY

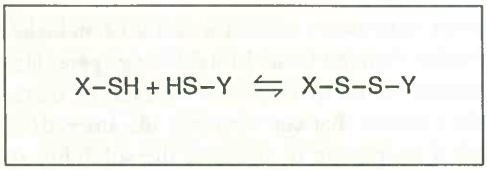


Figure 6.4.4 Equilibrium for the covalent interactions (formation of disulfide bonds) during the folding of a protein, X and Y are cysteine residues in the amino acid sequence of the folding protein that interact specifically in the folded structure.

At one extreme—where the conditions of folding (e.g., denaturant concentration and pH) lead to noncovalent interactions that are strong enough to stabilize the protein in a native-like form, with the native pairs of thiol groups forced close to one another (i.e., the equilibrium illustrated in Fig. 6.4.3 lies to the right)—very little oxidizing pressure will be needed to form the disulfide bonds. If, on the other hand, the noncovalent interactions are relatively weak (i.e., the equilibrium illustrated in Fig. 6.4.3 lies well over to the left and hence toward the expanded state of the protein), more energy—a stronger oxidation potential—will be required to drive the equilibrium to the native state (Gilbert, 1994).

To fold proteins containing disulfide bonds, it is necessary to pay attention to the chemistry of the covalent bonds involved in cysteine thiol oxidation. Two principal folding routes are possible—one utilizing air oxidation in the presence of trace heavy metals (Fig. 6.4.5) and the other involving formation of mixed-disulfide intermediates (Fig. 6.4.6).

The former route (Fig. 6.4.5), although cheaper when using large volumes, is less easy to optimize and control. The latter route (Fig. 6.4.6) utilizes a low-molecular-weight dithiol (R-S-S-R), usually glutathione. In practice, a redox pair of reduced glutathione (GSH) and oxidized glutathione (GSSG) is needed to create the appropriate oxidizing "pressure" that will allow the disulfides in the folding intermediates to make and break, so that the noncovalent interactions can shuffle the protein to its lowest free energy state. Because of the energy balances discussed above, the lowest free energy state will be expected to vary from protein to protein; hence the optimal redox conditions will also differ. However, a GSH:GSSG molar ratio of 10:1 (at a concentration of 2 to 5 mM GSH) gives maximal rates of folding to the native structure for a number of proteins. Because each step in the oxidation involves nucleophilic attack by an S^- group (see Fig. 6.4.6), the rate of disulfide formation is expected to increase within the range of pH 6 to 10. However, pH may also affect the stability of folding intermediates in an unpredictable way. In cases where the noncovalent interactions alone do not lead to a condensed intermediate (i.e., where the equilibrium illustrated in Fig. 6.4.3 lies to the left; see discussion above), addition of a stabilizing cosolvent such as phosphate or sulfate can drive the equilibrium in the direction of condensation, thereby offering the possibility of increasing the rate of disulfide for-

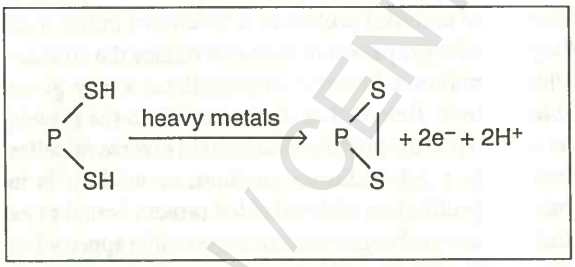


Figure 6.4.5 Formation of disulfide bond by air oxidation of thiol groups in the presence of trace heavy metals.

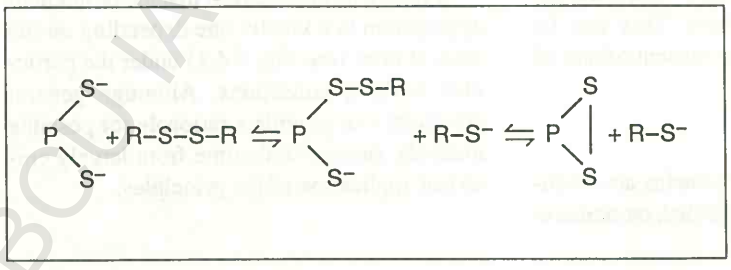


Figure 6.4.6 Formation of disulfide bond through mixed-disulfide intermediates.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

mation and interchange. Other factors (e.g., ionic strength and temperature) that can affect the stability of intermediates will also influence the kinetics and yield of folding.

Achieving an Acceptable Rate of Folding

Within the plateau region of the unfolding transition (see Fig. 6.4.2), the rate of folding will generally be in inverse proportion to the concentration of denaturant. The rate of folding will increase with temperature, up to that of thermal denaturation. For certain proteins whose folding is rate-limited by proline *cis-trans* isomerization (see Proteins Fold Along a Limited Number of Pathways, above), the rate can be increased by the addition of peptidyl prolyl *cis-trans* isomerase (Nall, 1994). However, when working with proteins lacking disulfide bonds, rate is not as important for practical purposes as yield.

The rate of folding for disulfide-bonded proteins in vitro can be very slow, with half-times of the order of minutes or hours, and conditions must be arranged to reduce the time required to realize an acceptable yield. The linkage between noncovalent and disulfide stability (see discussion of disulfide-bonded proteins under Stabilizing the Native Functional State) has important consequences in defining conditions for folding disulfide-bonded proteins. Certain intermediates in some folding pathways have been shown to possess nonnative disulfides (Gilbert, 1994). The energy barriers involved in making and breaking both covalent disulfide links and noncovalent interactions must therefore not be too high if such a pathway is to be followed. If, as a result of the redox conditions, the nonnative disulfides in the intermediates are so strong that they cannot be broken to allow them to shuffle, the native noncovalent interactions will not be able to form. It is necessary therefore to create a redox potential such that the equilibrium illustrated in Figure 6.4.4 is appropriately balanced. The converse of this requirement is that the equilibrium in Figure 6.4.3 must be appropriately balanced to allow the noncovalent interactions to shuffle and thereby allow the native disulfide pairs to form. This can be achieved by adjusting the concentrations of denaturant and cosolvent.

Limiting Aggregation

Aggregates of unfolded proteins are intrinsically more stable than the folded, monomeric state, as is shown by the formation of inclusion bodies. A major factor in stabilizing aggregates appears to be hydrophobic interaction; therefore factors that can decrease the strength of such interaction or increase the solubility of other parts of the protein chain will help relieve problems of low yield stemming from aggregation. The characteristic inverse dependence of stability on temperature in hydrophobic interactions suggests that lower temperatures should lessen aggregation. Nondenaturing concentrations of urea or guanidine·HCl have been also used with considerable success to lessen aggregation. These concentrations have usually been arrived at empirically, by dialyzing denaturant out of the denaturation mixture. Increasing the net charge by moving away from the isoelectric point and/or lowering the ionic strength may also lessen aggregation, other factors being equal.

Chaperone proteins, notably GroEL, have in certain cases been shown to dramatically increase the yield of refolding proteins in vitro (Goloubinoff et al., 1989). This multimeric protein, cylindrical in shape with a cylindrical core, will bind partially folded protein monomers, effectively removing them from free solution, and thereby protecting them against aggregation. The protein is then released with the cooperation of a second chaperone protein, GroES, under the influence of ATP.

The obvious factor governing aggregation is protein concentration (Hlodan et al., 1991; Thatcher and Hitchcock, 1994). Diluting the unfolded protein into larger volumes of buffer should increase the proportion of molecules that fold, although problems of recovery will also increase. Dropwise addition of the solution of unfolded protein to a volume of buffer over a longer period of time can reduce the concentration of "sticky" intermediates at any given time. Removal of denaturant from the protein while the protein is isolated in reverse micelles or a gel-exclusion medium, or while it is in equilibrium with unfolded protein bound to an ion-exchange column, are possible approaches to refolding at higher concentrations of protein (Thatcher and Hitchcock, 1994).

It should be emphasized that the problem of aggregation is a kinetic one depending on the ratio of rates (see Fig. 6.4.1) under the particular working conditions. Although general principles can provide a rationale for possible methods, success will come from largely empirical application of the principles.

BCU IASH CENTRAL UNIVERSITY LIBRARY

LITERATURE CITED

Anfinsen, C.B. 1967. The formation of the tertiary structure of proteins. *Harvey Lect.* 61:95-116.

Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. *Science* 181:223-230.

Craig, S., Schmeissner, U., Wingfield, P., and Pain R.H. 1987. Conformation, stability and folding of interleukin-1β. *Biochemistry* 26:3570-3576.

Gilbert, H.F. 1994. The formation of native disulfide bonds. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 109-111. Oxford University Press, Oxford.

Goloubinoff, P., Christeller, J.T., Gatenby, A.A., and Lorimer, G.H. 1989. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperone proteins and ATP. *Nature* 342:884-889.

Haase-Pettingell, C.A. and King, J. 1988. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tail spike maturation. *J. Biol. Chem.* 263:4977-4983.

Hlodan, R., Craig, S., and Pain, R.H. 1991. Protein folding and its implications for the production of recombinant proteins. *Biotechnol. & Genet. Eng. Rev.* 9:47-88.

Hlodan, R. and Hartl, F.U. 1994. How the protein folds in the cell. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 194-228. Oxford University Press, Oxford.

Lapanje, S., Skerjane, J., Glavnik, S., and Zibret, S. 1978. Thermodynamic studies of the interactions of guanidinium chloride and urea with some oligoglycines and oligolysines. *J. Chem. Thermodynam.* 10:425-433.

Lindsay, C.D. and Pain, R.H. 1991. Refolding and assembly of penicillin acylase, an enzyme composed of two polypeptide chains that result from proteolytic activation. *Biochemistry* 30:9034-9040.

Lomas, D.A., Evans, D.Ll., Stone, S.R., Chang, W.-S.W., and Carrell, R.W. 1993. Effect of the Z mutation on the physical and inhibitory properties of α_1-antitrypsin. *Biochemistry* 32:500-508.

Mitchinson, C. and Pain, R.H. 1985. Effects of sulfate and urea on the stability and reversible unfolding of β-lactamase from *Staphylococcus aureus*. *J. Mol. Biol.* 184:331-342.

Nall, B.T. 1994. Proline isomerization as a rate-limiting step. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 80-103. Oxford University Press, Oxford.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E., and Razgulaev, D.I. 1990. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 262:20-24.

Roder, H. and Elöve, G.A. 1994. Early stages of protein folding. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 37-40. Oxford University Press, Oxford.

Thatcher, D. and Hitchcock, A. 1994. Protein folding in biotechnology. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 242-250. Oxford University Press, Oxford.

Contributed by Roger H. Pain
Jozef Stefan Institute
Ljubljana, Slovenia

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Folding and Purification of Insoluble (Inclusion Body) Proteins from *Escherichia coli*

Heterologous expression of recombinant proteins in *E. coli* often results in the formation of insoluble and inactive protein aggregates, commonly referred to as inclusion bodies. To obtain the native (i.e., correctly folded) and hence active form of the protein from such aggregates, four steps are usually followed. In the first step, the bacterial cells are lysed and the aggregates collected by low-speed centrifugation (*UNITS 6.2 & 6.3*). In the second step (which is usually but not always included) the aggregates are preextracted to remove cell wall and outer membrane components, producing washed pellets (*UNIT 6.3*). In the third step, the aggregates are solubilized (or extracted) with strong protein denaturants—e.g., guanidine·HCl, 8 M urea, or organic acids at high concentrations (*UNIT 6.3*). In the fourth step, the solubilized, denatured proteins are folded with concomitant oxidation of reduced cysteine residues into the correct disulfide bonds to obtain the native protein. Basic Protocols 1, 2, and 3 in this unit feature three different approaches to the final step of protein folding and purification.

Basic Protocol 1 describes the extraction of bovine growth hormone (BGH) using guanidine·HCl, after which the solubilized protein is folded (before purification) in an "oxido-shuffling" buffer system to increase the rate of protein oxidation. A nondenaturing concentration of urea (a cosolvent; see *UNIT 6.4*) is added to the buffer used in the folding step to increase the yield of folded protein by minimizing the formation of aggregated, misfolded protein. The folded protein is then purified by chromatography.

Basic Protocol 2 describes folding and purification of the cytokine interleukin 2 (IL-2). This procedure differs from Basic Protocol 1 in that (1) acetic acid is used to solubilize (i.e., extract) the protein, (2) the acid-soluble protein is partially purified by gel filtration before folding, and (3) protein folding and oxidation are carried out with no additives; the protein (in acid solution) is simply dialyzed against water, thereby allowing the pH to slowly increase and air oxidation of the protein to occur. The folded protein is further purified by reversed-phase high-performance liquid chromatography (RP-HPLC). A support protocol is included for rapidly determining the amount of folded protein that contains the correct disulfide linkage pattern.

Basic Protocol 3 describes the folding and purification of a fusion protein containing the catalytic domain of the human immunodeficiency virus 1 (HIV-1) integrase. The fusion partner is a 20-amino-acid peptide containing a track of six histidine residues (a His tag; see *UNIT 5.1*). The protein is extracted with guanidine·HCl (as with BGH), then purified by gel filtration in guanidine·HCl (*UNIT 6.3*). The protein is further purified by metal-chelate affinity chromatography, which takes advantage of the His tag. Next, the purified protein is folded by very slow dilution into a folding buffer containing additives required to maintain the solubility of the folded protein. The His tag is removed by digestion with the protease thrombin, the thrombin is removed by affinity chromatography using an immobilized substrate, and finally the cleaved protein is run through a gel-filtration column to remove any aggregated protein.

BCU ALFRED TRIOLI UNIVERSITY LIBRARY

Copyright © 1995 by John Wiley & Sons, Inc.

BASIC PROTOCOL 1

FOLDING AND PURIFICATION OF BOVINE GROWTH HORMONE

The high-level expression of bovine growth hormone (BGH) in *E. coli* results in the formation of highly aggregated protein (inclusion bodies). To restore the BGH to its native, functional state, the *E. coli* cells are first lysed with the French press and the inclusion bodies are partially purified by selective extraction of soluble and particulate *E. coli* contaminants. The inclusion bodies are solubilized with guanidine·HCl to give protein that is unfolded and reduced (i.e., with four free sulfhydryl residues). The protein is then simultaneously folded and oxidized by the slow removal of protein denaturant, resulting in the formation of two disulfide bonds from the sulfhydryl residues. The solubility of the protein during folding is maintained using a cosolvent (urea) and the oxidation of cysteine residues is catalyzed by a mixture of reduced and oxidized glutathione (GSH/GSSG; see also *APPENDIX 3A*). The folded, oxidized protein is finally purified to homogeneity by anion exchange and gel-filtration chromatography.

Materials

E. coli cells expressing BGH: ~50 g wet weight from a 1.5-liter fermentation (*UNIT 5.3*), and stored as a flattened paste in a sealed polyethylene bag at –80°C
BGH break buffer A (see recipe), 4°C
BGH break buffer B (see recipe), 4°C
BGH wash buffer A (see recipe), 4°C
BGH wash buffer B (see recipe), 4°C
BGH extraction buffer (see recipe), 4°C
BGH folding buffer A (see recipe), 4°C
BGH folding buffer B (see recipe), 4°C
2 M HCl
BGH column buffer A (see recipe), 4°C
DEAE Sepharose CL-4B ion-exchange resin (Pharmacia Biotech)
Sephadex G-100 gel filtration resin (Pharmacia Biotech)
BGH column buffer B (see recipe), 4°C

Blender (e.g., Waring; ≥1 liter capacity)
Tissue homogenizer (e.g., Polytron, Brinkmann)
Beckman J2-21M centrifuge and JA-20 rotor (or equivalent)
Sonicator, ≥400 W, with sound enclosure (Branson or equivalent)
0.7-µm glass-microfiber filters, 4.7-cm diameter (Whatman GF/F) with vacuum filtration apparatus (optional)
Spectra/Por 1 dialysis tubing, 40-mm diameter (MWCO 6000 to 8000; Spectrum)
5 × 50–cm and 5 × 100–cm glass chromatography columns with adjustable flow adaptors (Pharmacia Biotech)
Stirred cell with Diaflo PM 10 ultrafiltration membrane (Amicon)
Millex-GV 0.22-µm filter units (Millipore)

Additional reagents and equipment for cell breakage using a French press (*UNIT 6.2*), dialysis (*APPENDIX 3B*), ion-exchange chromatography (*UNIT 8.2*), gel-filtration chromatography (*UNIT 8.3*), and SDS-PAGE (*UNIT 10.1*)

NOTE: All steps are carried at 4°C unless otherwise stated.

Break cells

1. Resuspend ~ 50 g (wet weight) of *E. coli* cells expressing BGH in 250 ml BGH break buffer A using a blender.

   The lysozyme in BGH break buffer A aids in removal of the peptidoglycan and outer membrane protein contaminants (see UNIT 6.1).

PCU IASI CENTRAL UNIVERSITY LIBRARY

2. Stir cell suspension 30 min at 20° to 25°C using a heavy-duty magnetic stirrer. To reduce viscosity of the suspension, homogenize with tissue homogenizer until no clumps are detected, then centrifuge 35 min at 20,000 × *g* (13,500 rpm in JA-20 rotor), 4°C.

3. Resuspend pellet in 150 ml BGH break buffer B using tissue homogenizer. Break cells using French press as described in *UNIT 6.2*, Basic Protocol 1, steps 1 to 5.

The viscosity can also be reduced by digestion with DNase and RNase as described in UNIT 6.2, in which case the EDTA in BGH break buffer B should be replaced with 5 mM $MgCl_2$.

Prepare and extract washed pellet

4. Centrifuge suspension 40 min at 20,000 × *g*. Resuspend pellet in 250 ml BGH wash buffer A using the tissue homogenizer, then centrifuge 30 min at 20,000 × *g*. Repeat centrifugation and resuspension once using BGH wash buffer A and again using BGH wash buffer B. Discard the cloudy supernatants and save the washed pellet.

If not to be used immediately, the washed pellet may be stored at –80°C.

5. Resuspend washed pellet in 150 ml BGH extraction buffer. Sonicate briefly to completely disperse protein, then clarify solution by centrifuging 30 min at 20,000 × *g* or by filtering through a 0.7-µm glass-microfiber filter under vacuum.

Generally, when protein is extracted with guanidine·HCl, the reductant dithiothreitol is included in the buffer system to ensure that the protein is fully reduced prior to folding. In this procedure it was established that >80% of the protein in the inclusion bodies was already in the reduced state (P.T.W., unpub. observ.). In the extraction buffer used here, therefore, only the relatively weak reductant glutathione is included. This compound serves to maintain thiols in the reduced state, and is a component of the redox buffer that will be used later to oxidize the protein. It cannot, however, be assumed that other inclusion-body proteins contain only free thiols.

Fold protein

6. Dilute the clear amber-colored solution with an equal volume of BGH folding buffer A and pour into prewashed dialysis tubing.

Use two or three pieces of tubing and fill dialysis bags only three-quarters of the way to allow for any volume increase during dialysis.

The BGH concentration from step 5 should be 2 to 4 mg/ml, and after dilution should not exceed 1.0 to 2.0 mg/ml. If higher concentrations are found, the solution must be diluted further with folding buffer A. BGH concentration can be estimated by diluting the sample with 4 M guanidine·HCl in water and measuring the absorbance at 280 nm and 260 nm in a cell with a 1-cm path length. The total protein concentration (mg/ml) is estimated as 1.55 A_{280} – 0.76 A_{260} nm (Stoscheck, 1990). The BGH content may either be assumed to be 60% of the total protein or, more accurately, be estimated by performing SDS-PAGE and densitometry using the washed pellet from step 4.

Urea is included in folding buffer A as a cosolvent to maintain solubility of the protein during refolding. Removal of guanidine·HCl by dialysis or dilution results in precipitation if no cosolvent is used. The urea concentration chosen (in this case 4 M) should be low enough to allow the native structure to form (see UNIT 6.4). Urea unfolding/folding profiles for BGH (i.e., equilibrium-denaturation curves in which protein conformation is measured as a function of denaturant concentration; see UNIT 6.4) were available in the literature prior to development of this method (Edelhoch and Burger, 1966). The urea concentration that induces protein unfolding can be determined rapidly by urea gradient electrophoresis (Goldenberg, 1989).

7. Dialyze solution 12 to 16 hr (or overnight) against 5 liters of BGH folding buffer A, then dialyze an additional 6 to 8 hr against 5 liters of BGH folding buffer B.

The second dialysis can be continued overnight if necessary.

BOURNEMOUTH UNIVERSITY LIBRARY

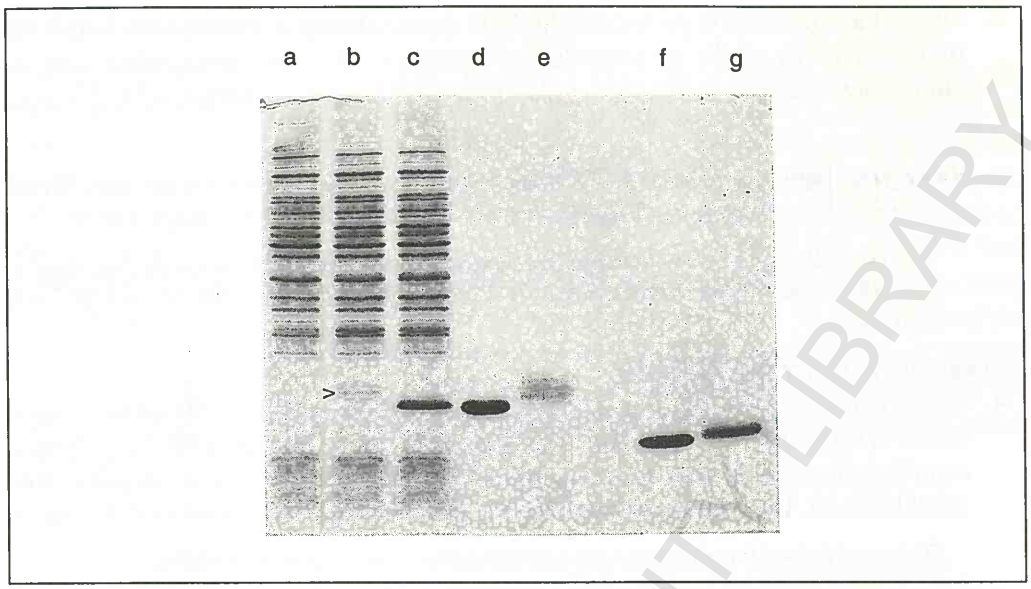


Figure 6.5.1 SDS-PAGE of bovine growth hormones on 12.5% polyacrylamide gel. Lane A, BGH-expressing *E. coli* cells minus the expression vector; lanes B and C, BGH-expressing *E. coli* cells with ΔA-4 and A-9 BGH expression vectors, respectively; lane D, purified recombinant A-9 BGH; lane E, BGH purified from pituitary (supplied by A.F. Parlow, UCLA); lane F, purified recombinant A-9 BGH with no reductant in sample buffer; lane G, BGH purified from pituitary with no reductant in sample buffer. In lane E (pituitary BGH), the two bands correspond to full-length protein and protein truncated at the N-terminus by 4 residues. It can be seen that the bottom band has the same mobility as *E. coli* extracts containing the ΔA-4 BGH construct. In lane G, it it may be noted that the two bands are not resolved under nonreducing conditions.

Oxidation of the protein during dialysis can be monitored by SDS-PAGE (UNIT 10.1). SDS-denatured oxidized BGH is more compact than SDS-denatured reduced BGH and thus migrates faster as a result of the lower apparent molecular weight (i.e., 18 kDa for the oxidized form versus 22 kDa for the reduced form). SDS-PAGE band patterns of oxidized and reduced BGH are shown in Figure 6.5.1. Any free thiol groups in the sample are quenched by addition of 20 mM iodoacetamide. SDS sample buffer (UNIT 10.1) minus reductant is then added. The pH of the SDS-treated sample may have to be readjusted with dilute alkali. The proportion of oxidized versus reduced protein is finally determined by densitometry of the Coomassie blue–stained gel. It should be noted that this approach does not prove whether or not the correct disulfide bond(s) have been formed. The shift in mobility upon reduction occurs because of the formation of the disulfide bond linking Cys-51 to Cys-163. BGH in which the second disulfide bond (linking Cys-180 to Cys-188) has been selectively reduced (Graf et al., 1975) still exhibits the gel shift. Despite these potential pitfalls, the gel method is useful and correlates with other approaches—e.g., direct monitoring of disulfide formation using 2-nitro-5-thiosulfobenzoate in the presence of sodium sulfite after the free sulfhydryls have been quenched and the buffer components removed (Thannhauser et al., 1984). Sulfhydryl groups can be assayed using Ellmans reagent (Riddles et al., 1979).

8. Centrifuge the slightly cloudy solution 30 min at 20,000 × *g*. Discard pellet and adjust supernatant to pH 9.0 with 2 M HCl.

The volume of the supernatant should be 300 to 320 ml.

Purify protein by ion-exchange and gel filtration chromatography

9. Pack a 5 × 50–cm column with DEAE Sepharose CL-4B to give a bed height of ~10 cm and equilibrate with BGH column buffer A. Apply the solution from step 8 at a flow rate of 60 ml/hr and begin collecting 15-ml fractions.

See UNIT 8.2 for further information on preparing an ion-exchange column.

10. Elute with BGH column buffer A at the same flow rate and continue fraction collection until A_{280} or A_{260} of eluant approaches a baseline value. Assay fractions by SDS-PAGE and pool fractions containing BGH.

Under the ion-exchange conditions used, BGH does not bind to the matrix and is located in the flowthrough fractions. The more acidic E. coli contaminants bind tightly to the top portion of the column, which turns brown. Some of the earlier flowthrough fractions may be slightly contaminated with aggregated protein that separates from soluble protein as a result of the gel-filtration effect of the matrix. A BGH concentration of 0.2 to 0.3 mg/ml in ~400 to 450 ml of pooled eluant should be obtained. This should exhibit a single band on the SDS-PAGE gel.

Pooled fractions can be stored 1 to 2 days at 4°C and for longer periods at –80°C.

11. Concentrate pooled BGH-containing fractions to 50 to 60 ml (i.e., to ~2 to 3 mg/ml protein concentration) using a stirred cell with Diaflo PM 10 ultrafiltration membrane. Centrifuge filtrate 15 min at 20,000 × *g* to remove any precipitated protein.

12. Pack a 5 × 100–cm column with Sephadex G-100 to a bed height of 95 cm and equilibrate with BGH column buffer B. Apply the filtrate from step 11 at a flow rate of 40 ml/hr and begin collecting 15-ml fractions.

See UNIT 8.3 for additional information on preparing gel-filtration columns. If a 5 × 100–cm column is not available, a 2.5 × 100–cm column can be used. Depending on the amount of sample, the smaller-diameter column may be run several times, loading 18 to 20 ml of sample per run.

13. Elute with BGH column buffer B at the same flow rate and continue fraction collection until A_{280} or A_{260} of eluant approaches a baseline value. Assay fractions by SDS-PAGE and pool fractions containing BGH.

The protein elutes in a single but slightly asymmetrical peak.

The gel-filtration elution peak has a sharp front edge, whereas the descending portion is more diffuse and trailing. This elution behavior results from the fact that BGH is a rapidly associating/dissociating monomer/dimer system (with a K_d of ~0.8 to 1.0 × 10^{-5} M; see Ackers, 1970, for discussion of theory). The apparent molecular weight estimated by gel filtration is 30 to 35 kDa (with a monomer mass of ~22 kDa).

The slightly alkaline pH of the column buffer and folding buffers is required to maintain solubility of the protein.

14. Repeat step 11, then filter sterilize filtrate using a Millex-GV 0.22-µm filter unit. Store purified BGH in aliquots at –80°C.

For long-term storage the protein can be lyophilized. If this is to be done, the pooled fractions from step 10 should be directly dialyzed (APPENDIX 3B) against 0.1 M ammonium bicarbonate (pH 9.2 to 9.4, adjusted with ammonium hydroxide), and the dialysate filter-sterilized and freeze-dried. If an essentially salt-free protein is required, two cycles of freeze-drying should be performed and the protein reconstituted with water alone after the first drying cycle.

The concentration of the purified BGH can be conveniently determined by UV absorbance measurement—i.e., 1 mg/ml native BGH has an A_{280} of 0.7 in a 1-cm quartz cuvette. Protein concentration may also be estimated in crude extracts using the Bio-Rad Protein Assay Kit (based on the Bradford method) and in partially purified fractions as described in the annotation to step 6. Biological activity of the protein is measured as described in Wingfield et al. (1987a).

BCU LIBRARY - FAITRN - UNIVERSITY LIBRARY

BASIC PROTOCOL 2

FOLDING AND PURIFICATION OF HUMAN INTERLEUKIN 2

Human interleukin 2 (hIL-2) is a hydrophobic, acid-stable polypeptide with a molecular weight of 15 kDa. Recombinant hIL-2 expressed in *E. coli* is deposited in inclusion bodies. Extraction of the inclusion bodies with denaturants followed by folding in aqueous solution often results in formation of insoluble hIL-2 aggregates. In this protocol, inclusion bodies containing recombinant hIL-2 are first extracted with acetic acid and the soluble monomeric hIL-2 is separated from other hIL-2 aggregates by gel-filtration chromatography. The soluble protein is then folded and oxidized by dialysis against water. Correctly folded hIL-2 is finally separated from unfolded hIL-2 and other *E. coli* contaminants by reversed-phase high-performance liquid chromatography (RP-HPLC).

Materials

hIL-2 break buffer (see recipe), 4°C
E. coli cells expressing hIL-2: ~20 g wet weight from a 3-liter fermentation (*UNIT 5.3*), and stored as a flattened paste in a sealed polyethylene bag at –80°C
Sucrose
Lysozyme (Worthington)
hIL-2 wash buffer: 0.75 M guanidine·HCl/1% (w/v) Tween 20 (prepare immediately before use), 4°C
PBS (*APPENDIX 2E*), 4°C
10% and 20% (v/v) acetic acid, 4°C (prepare fresh from glacial acetic acid)
Sephadex G-100 gel-filtration resin (Pharmacia Biotech)
Acetonitrile (HPLC grade)
Trifluoroacetic acid (TFA; HPLC grade)
7 × 250–mm 300-Å octyl Aquapore RP-300 semiprep column (Brownlee column; Thomson Instrument)
RP-HPLC solvent A (see recipe), room temperature
RP-HPLC solvent B (see recipe), room temperature
25 mM acetic acid, 4°C

Tissue homogenizer (e.g., Polytron, Brinkmann)
30°C water bath
Sorvall RC-5C centrifuge with SS-34 rotor (or equivalent)
2.6 × 100–cm glass chromatography column
Spectra/Por 3 dialysis tubing, 11.5- and 45-mm diameters (MWCO 3500; Spectrum)
Sterivex-GS 0.22-μm filter units (Millipore)
HPLC system with pumps, UV detector, and fraction collector (Waters)

Additional reagents and equipment for cell breakage using a French press (*UNIT 6.2*), gel-filtration chromatography (*UNIT 8.3*), and dialysis (*APPENDIX 3B*)

Break cells

1. Add 60 ml hIL-2 break buffer to 20 g (wet weight) of *E. coli* cells expressing hIL-2 in a 250-ml glass beaker. Mix well using tissue homogenizer.

   UNIT6.2 describes use of a tissue homogenizer to resuspend E. coli cells.

2. Add 21 g sucrose and mix well with cell paste using tissue homogenizer. Add 34 mg lysozyme and mix again using tissue homogenizer. Incubate 30 min in a 30°C water bath, then dilute with 100 ml hIL-2 break buffer and cool on ice.

3. Break cells by passing suspension through a French press twice as described in *UNIT 6.2*. Centrifuge 30 min at 13,000 × *g* (10,400 rpm in SS-34 rotor), 4°C. Save pellet.

   The wet weight of the pellet should be ~1.5 g.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

4. Wash pellet by resuspending in 30 ml hIL-2 wash buffer and centrifuging 30 min at 13,000 × *g*. Repeat wash once with hIL-2 wash buffer and again using 30 ml PBS in place of the wash buffer.

Washed pellets should be used immediately or stored at –70°C until required.

Extract protein with acid and prepare gel-filtration column

5. Add 12.5 ml of 20% acetic acid to the washed pellet and mix using tissue homogenizer. Centrifuge 30 min at 13,000 × *g*, 4°C, and save the supernatant.

6. Pack a 2.6 × 100–cm column with Sephadex G-100 to give a bed height of 95 cm and equilibrate with 10% acetic acid.

For convenience the column may be packed the day before use.

7. Repeat step 5 and pool the two clear supernatants (total volume should be ~25 ml).

Isolate hIL-2 monomer from solubilized protein by gel filtration

8. Immediately apply pooled supernatants to column prepared in step 6. Perform gel filtration chromatography at 4°C as described in *UNIT 8.3* and construct a chromatogram, pooling the fractions that make up the third peak.

The solution should be applied to the column immediately after extraction to prevent any covalent modification of the protein.

The column eluate usually shows three major peaks in A_{280}. The first peak (in the void volume) contains aggregates, the second peak contains dimeric hIL-2, and the third peak contains monomeric hIL-2 (S.M.L., unpub. observ.). A T cell proliferation assay shows that the third peak contains the highest biological activity of hIL-2 (Liang et al., 1986; Bottomly et al., 1991).

9. Using 45-mm diameter Spectra/Por 3 dialysis tubing (MWCO 3500), dialyze the pooled fractions (~60 ml) making up the third peak at 4°C overnight against 5 liters of Milli-Q water, then for an additional 3 to 4 hr against 5 liters of fresh Milli-Q water.

The pooled eluate fractions should be extensively dialyzed against water to remove acetic acid. The dialysate should have a volume of ~60 ml with an hIL-2 concentration of 0.2 mg/ml. After dialysis, the hIL-2 monomer can be stored at –20°C until required.

As the acetic acid is removed; sample pH slowly increases, thereby allowing hIL-2 thiols to form disulfide bonds.

Purify protein by RP-HPLC

10. Filter the dialysate through a 0.22-µm filter unit, then add 30 ml acetonitrile (33% v/v final) and 0.09 ml TFA (0.1% v/v final).

11. Prepare and test the HPLC system according to manufacturer's instructions. Place a 300-Å octyl Aquapore RP-300 reversed-phase chromatography column in line.

Other reversed-phase columns; e.g., the 2.2 × 15–cm TMS-250 (TosoHaas) can also be used (Kato, 1985).

12. Fill the reservoirs of the HPLC system gradient maker with RP-HPLC solvent A and RP-HPLC solvent B. Carry out a blank run (at room temperature) with the following gradient program:

0 min: 0% solvent B/100% solvent A
10 min: 0% solvent B/100% solvent A
60 min: 50% solvent B/50% solvent A
90 min: 50% solvent B/50% solvent A
150 min: 70% solvent B/30% solvent A

BCU/ASU/CENTRAL UNIVERSITY LIBRARY

160 min: 70% solvent B/30% solvent A
170 min: 100% solvent B/0% solvent A
180 min: 100% solvent B/0% solvent A
190 min: 0% solvent B/100% solvent A
210 min: 0% solvent B/100% solvent A.

Repeat blank run as necessary to ensure that the baseline profile has stabilized.

13. Pump sample from step 10 into column at a flow rate of 4 ml/min. Run the gradient program described in step 12 at a flow rate of 1 ml/min at room temperature. Collect 8-ml fractions at a rate of 1 ml/min and pool the fractions making up the correctly folded hIL-2 peak, which elutes at ~70 min.

Care should be taken to avoid introducing air bubbles while pumping the sample into the column, as these will generate false peaks in the chromatogram. If the peaks in the HPLC profile are too broad, the gradient program should be varied to improve the separation (see Chapter 8).

14. Using 11.5-mm diameter Spectra/Por 3 dialysis tubing (MWCO 3500), dialyze the pooled hIL-2 fractions against Milli-Q water or 25 mM acetic acid at 4°C. Filter the purified protein solution using a 0.22-μm filter unit and store at –20°C or below.

SUPPORT PROTOCOL

RESOLUTION OF NATIVE AND MISFOLDED FORMS OF hIL-2 BY RP-HPLC

Human interleukin 2 (hIL-2) contains three cysteine residues—at positions 58, 105, and 125. In the correctly folded (native) protein, the cysteines at positions 58 and 105 form a disulfide bond that is important for biological activity of the protein (Robb et al., 1984; Wang et al., 1984; Liang et al., 1985).

As hIL-2 contains three cysteines, three possible intramolecular disulfide–linked forms can exist, only one of which is the native form. RP-HPLC can separate these isomers, and is used to verify that the purified IL-2 contains the native disulfide.

Additional Materials (also see Basic Protocol 2)

0.46 × 10–cm SynChropak RP-P C18 column (Thomson Instrument)
RP-HPLC solvent C (see recipe)
25 mM acetic acid
Purified folded hIL-2 solution (see Basic Protocol 2)

1. Prepare and test HPLC system according to manufacturer's instructions. Place a 0.46 × 10–cm SynChropak RP-P C18 column in line.
2. Fill the reservoirs of the HPLC system gradient maker with RP-HPLC solvent A and RP-HPLC solvent C. Apply 20 μl of 25 mM acetic acid as a blank sample. Perform a blank run at room temperature using a gradient program starting with 100% solvent A for 5 min, followed by a linear gradient starting at 100% solvent A/0% solvent C and ending at 0% solvent A/100% solvent C, at a flow rate of 1 ml/min. Repeat as necessary to check that the baseline profile is reproducible.
3. Prepare a solution of ~5 mg/ml purified folded hIL-2 solution in 25 mM acetic acid and filter through a 0.22-μm filter.
4. Apply 20 μl of purified folded hIL-2 solution in 25 mM acetic acid to the column and carry out RP-HPLC at room temperature using the gradient program described in step 2. Collect the fractions containing the properly folded hIL-2.

PCU IASI CENTRAL UNIVERSITY LIBRARY

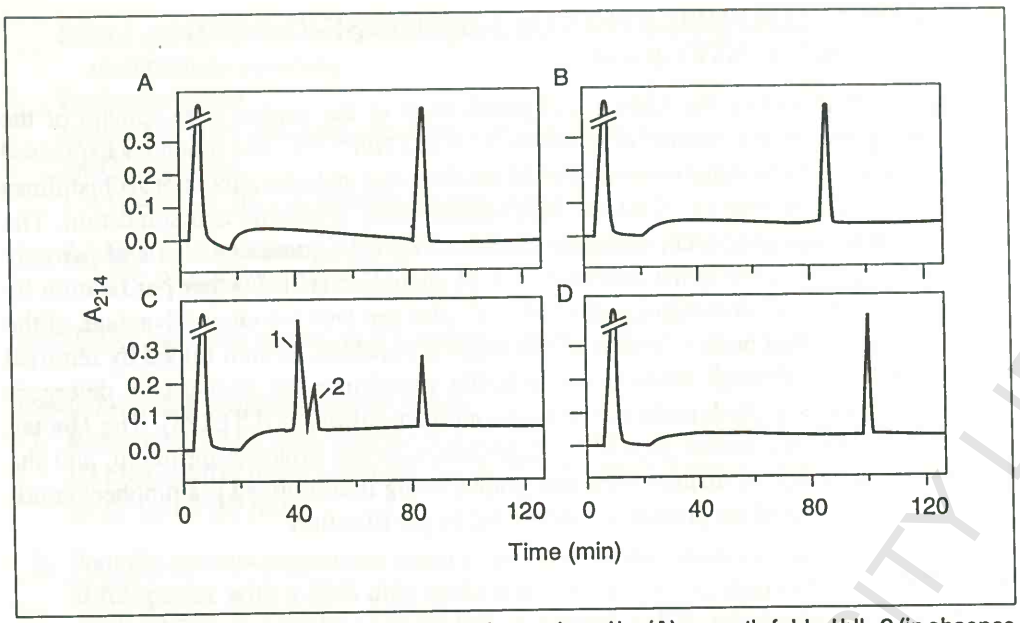


Figure 6.5.2 Chromatograms illustrating peaks produced by (**A**) correctly folded hIL-2 (in absence of denaturant); (**B**) correctly folded hIL-2 (in presence of denaturant but at pH 3.5, which is too low for disulfide-bond exchange); (**C**) scrambled hIL-2 isomers (resulting from denaturant treatment at pH 8.5); (**D**) unfolded hIL-2 (resulting from denaturant/reductant treatment).

Properly folded hIL-2 usually elutes at ~54% acetonitrile and unfolded hIL-2, which has no intramolecular disulfide bond, elutes at 57% acetonitrile.

In the presence of denaturants such as guanidine·HCl, hIL-2 rapidly scrambles into a mixture of three disulfide-linked isomers. The incorrectly folded hIL-2 elutes at a lower acetonitrile concentration (<54%) as shown in Figure 6.5.2. For each separation in Figure 6.5.2, two Vydac C4 columns were connected in series and run at 40°C with a 1 ml/min flow rate. The sample was eluted with an acetonitrile gradient that increased from 0% to 44% in 15 min, and then to 64% in 120 min. In the separation represented by the chromatogram in panel A, 100 µg hIL-2 in 40 µl of 50 mM acetic acid was dissolved in 400 µl of 20% formic acid and chromatographed. The elution peak (at ~57.3% acetonitrile) represents native hIL-2. In panel B, hIL-2 was added to 300 µl of 6 M guanidine·HCl/0.2 M acetic acid, pH 3.5, and incubated 10 min at room temperature. 200 µl of 20% formic acid was added and the sample was chromatographed. Note that the chromatogram has not changed, as the pH is below the point where isomerization (disulfide-bond exchange) occurs. In panel C, hIL-2 was added to 300 µl of 6 M guanidine·HCl/0.05 M Tris·Cl, pH 8.5, incubated 5 min at room temperature, then quenched with formic acid as in the separation represented by panel B. Peak 1 represents the scrambled isomer with a disulfide bond between Cys-105 and Cys-125, whereas peak 2 is probably the scrambled isomer with a disulfide bond between Cys-58 and Cys-125. Finally, in panel D, hIL-2 was added to 6 M guanidine·Cl as in panel C. 50 mM DTT was then added and the sample was incubated 40 min at room temperature, then quenched with formic acid as in panel B. The peak (at ~59.4% acetonitrile) represents completely reduced (unfolded) hIL-2. (see also Browning et al., 1986).

IMPORTANT NOTE: *If the pressure in the HPLC column exceeds 4000 psi, the column should be cleaned.*

B. C. U. „M. ... IAŞI

BCU IAŞI - CENTRAL UNIVERSITY LIBRARY

FOLDING AND PURIFICATION OF A HISTIDINE-TAGGED PROTEIN: HIV-1 INTEGRASE

This protocol describes the folding and purification of the central core domain of the HIV-1 integrase, which consists of residues 50 to 212 (IN^{50-212}). The protein is expressed in *E. coli* with an N-terminal extension of 20 residues that includes a track of six histidines (i.e., a His tag; see *UNIT 5.1*). Cell are first broken using lysozyme and sonication. The protein is then extracted from inclusion bodies using 8 M guanidine·HCl and partially purified by gel filtration in the presence of 4 M guanidine·HCl. Further purification by metal-chelate affinity chromatography (MCAC; also see *UNIT 9.4*) takes advantage of the His tag. The purified protein, which at this stage is unfolded, is then folded by removal of guanidine·HCl through dilution into a buffer containing the zwitterionic detergent 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS). The His tag is removed from the folded protein by digestion with the protease thrombin, and the thrombin is removed by affinity chromatography using immobilized *p*-aminobenzamidine. A final cleanup of the protein is carried out by gel filtration.

Materials

E. coli cells expressing HIV-1 integrase: 100 g wet weight from a 3.5-liter fermentation (*UNIT 5.3*), stored as a flattened paste in a sealed polyethylene bag at −80°C
IN break buffer (see recipe), 4°C
Suspension buffer: IN break buffer (see recipe) prepared without lysozyme, 4°C
IN extraction buffer (see recipe), 4°C
6 × 60–cm Superdex 200 prep grade prepacked gel-filtration column (Pharmacia Biotech)
IN column buffer A (see recipe), 4°C
Ni-NTA-Sepharose CL-6B MCAC resin (Qiagen)
IN column buffer B (see recipe), 4°C
IN column buffer C (see recipe), 4°C
IN column buffer D (see recipe), 4°C
6 × 60–cm Superdex 75 prep grade prepacked gel-filtration column (Pharmacia Biotech)
IN column buffer E (see recipe), 4°C
4 M guanidine·HCl/5 mM DTT (4°C)
IN folding buffer (see recipe), 4°C
50 mM Tris·Cl (pH 7.5 at 4°C)/10 mM CHAPS, 4°C
2000 NIH unit/ml thrombin (see recipe)
IN column buffer F (see recipe), 4°C
9 M urea, 4°C
IN column buffer G (see recipe), 4°C
p-aminobenzamidine immobilized on Sepharose 6B (Pierce, Pharmacia Biotech, or Sigma)
0.1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF; Boehringer Mannheim or ICN Biomedicals)

Blender (e.g., Waring; ≥1 liter capacity)
1-liter steel beaker
Sonicator, ≥400 W, with sound enclosure (Branson or equivalent)
Centrifuge with Beckman J-14 rotor (or equivalent)
Tissue homogenizer (e.g., Polytron, Brinkmann)
Ultracentrifuge with Beckman 45Ti rotor (or eqivalent)
5 × 50–cm chromatography column with adjustable flow adaptors

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Stirred cells (400-ml and 2-liter capacities) with Diaflo PM 10 ultrafiltration membranes (Amicon)
Peristaltic pump
Millex-GV 0.22-μm pore size filter units (Millipore)
28°C water bath
1 × 10– to 1 × 20–cm chromatography column

Additional reagents and equipment for gel-filtration chromatography (*UNIT 8.3*) and SDS-PAGE (*UNIT 10.1*)

NOTE: All steps are carried at 4°C unless otherwise stated.

Break cells and prepare washed pellets

1. Resuspend 100 g (wet weight) of *E. coli* cells expressing HIV-1 integrase in 400 ml IN break buffer using a blender. Stir with a magnetic stirrer at room temperature (20° to 25°C) for 30 min.

2. Pour the viscous suspension into a 1-liter steel beaker in an ice bath. Sonicate 20 min at full power with a 60% duty cycle (i.e., on for 0.6 sec, then off for 0.4 sec), with stirring.

   Be sure that the temperature does not rise above 20°C.

3. Centrifuge 1 hr at 30,000 × *g* (14,000 rpm in J-14 rotor), 4°C. Resuspend pellet with 500 ml suspension buffer and centrifuge 1 hr at 30,000 × *g*.

Extract washed pellet and purify by gel filtration

4. Dissolve pellet in 40 ml IN extraction buffer using tissue homogenizer. Centrifuge 45 min at 100,000 × *g* (35,000 rpm in 45Ti rotor), 4°C.

5. Apply the clear supernatant to a 6 × 60–cm Superdex 200 gel-filtration column (*UNIT 8.3*) equilibrated with IN column buffer A. Elute at 300 ml/hr using IN column buffer A, collecting 15-ml fractions. Assay fractions by SDS-PAGE (*UNIT 10.1*) and pool those containing $IN^{50\text{-}212}$.

   $IN^{50\text{-}212}$ should be the major band visible on SDS-PAGE after gel filtration, representing 20% to 50% of the total protein.

Purify $IN^{50\text{-}212}$ by MCAC

6. Apply the $IN^{50\text{-}212}$ sample to a 5 × 50–cm column that has been packed with Ni-NTA-Sepharose CL-6B MCAC resin to a bed height of 5 cm and equilibrated with IN column buffer B at a flow rate of 300 ml/hr. Continue washing column with buffer until absorbance approaches the baseline.

7. Continue washing the column with 5 to 10 column volumes (490 to 980 ml) of IN column buffer C.

   The absorbance of the column eluate should be close to baseline before the sample is applied.

   Histidine is bound to the Ni-NTA resin by coordination bonding of the nonprotonated nitrogen atom of the imidazole side chain to Ni^{2+} ions. Because the pH of the IN column buffer C (6.4) is close to the average pK_a of histidine (6.0 to 7.0), only proteins that contain a track of six histidines (i.e., the His tag) are bound strongly enough to be retained. The sample applied to the column should not contain EDTA, and if reductant is required 1 mM 2-mercaptoethanol should be used rather than dithiothreitol.

   The Ni-NTA matrix is reported to have a binding capacity of ~5 to 10 mg protein per ml resin.

BOSTON COLLEGE CENTRAL UNIVERSITY LIBRARY

8. Elute bound protein with a linear pH gradient in a total volume of 980 ml (10 column volumes) beginning at pH 6.4 and ending at pH 4.5: make the gradient by placing 490 ml IN column buffer C (pH 6.4) and 490 ml of IN column buffer D (pH 4.5) in the appropriate reservoirs of the gradient maker. Elute column at ~300 ml/hr, collect 10-ml fractions and pool those making up the major peak (containing $IN^{50\text{-}212}$; eluted at ~pH 5.5).

   After elution, the column can be cleaned by washing with 0.2 M acetic acid/6 M guanidine·HCl, then reequilibrated with IN column buffer B. Under the conditions described, the column can be used two or three times before it turns from light blue-green to brownish-gray, indicating that regeneration is required. The column can be regenerated by passing 2 column volumes of 100 mM EDTA to strip Ni from the resin, washing the column sequentially with 2 column volumes each of 0.1 M NaOH and 6 M guanidine·HCl, and finally passing 2 column volumes of 10 mM $NiSO_4$ through the column to charge the resin with Ni. The column is then washed with water and equilibrated with IN column buffer B. Further details on use of the Ni-NTA resin are given in the manufacturer's literature (Qiagen, 1992) and in UNIT 9.2.

9. Concentrate pooled $IN^{50\text{-}212}$-containing fractions to ~40 ml (i.e., ~8 to 10 mg/ml protein) using a stirred cell with Diaflo PM 10 ultrafiltration membrane.

10. Apply concentrated sample to a 6 × 60–cm Superdex 75 gel-filtration column equilibrated with IN column buffer E. Pool the peak fractions.

    Reductants (either dithiothreitol or 2-mercaptoethanol) are included in the buffers to maintain the $IN^{50\text{-}212}$ cysteine thiol groups in the reduced state. It should be noted that $IN^{50\text{-}212}$ contains three cysteine residues that do not participate in disulfide bond formation.

    Pooled fractions may be stored in 10 to 20 ml aliquots at –80°C until ready to use.

    *This gel-filtration step removes minor protein contaminants and metal ions leached from the MCAC column. It also exchanges the protein into fresh buffer containing DTT. Alternatively, the pooled fractions from step 8 can be concentrated by ultrafiltration to ≤1 mg/ml protein, then dialyzed (*APPENDIX 3B*) against IN column buffer E.*

Fold protein by dilution

11. Dilute pooled fractions from step 10 to ~1 mg/ml protein using 4 M guanidine·HCl/5 mM DTT. Pump 30 ml (or other volume as required) of the diluted solution into 1 liter IN folding buffer using a peristaltic pump at a flow rate of ~3 ml/hr while stirring gently with a magnetic stirrer (see Fig. 6.5.3).

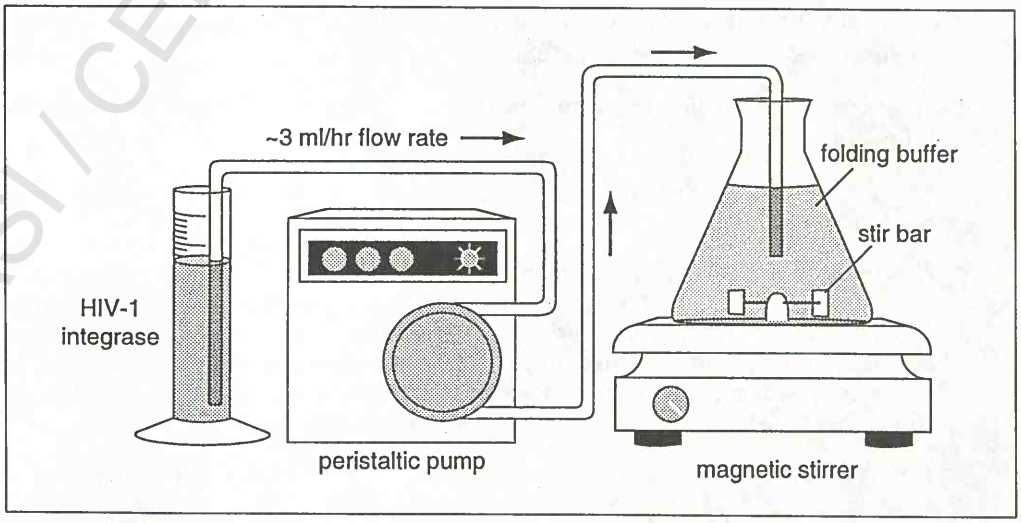


Figure 6.5.3 Setup for folding of HIV-1 integrase by dilution into buffer.

ECU IAST / CENTRAL UNIVERSITY LIBRARY

After dilution, the final protein concentration will be ~30 μg/ml.

Protein concentration can be estimated using A_{280} = 1.46 for 1 mg/ml folded IN^{50-212} (still containing the His tag) measured in a cuvette with a 1-cm path length.. If A_{280} is >2.0, dilute with 4 M guanidine·HCl in water.

The scale of operation at this step is limited by the large dilution of the protein. This step can be scaled up if equipment suitable for concentration of large volumes is available (see annotation to step 12). Otherwise, the step can be repeated if more protein is needed.

Protein folding is usually carried out at a low concentration to avoid aggregation, which is thought to arise from partially folded protein in the folding pathway (i.e., folding intermediates; Goldberg et al., 1991; Jaenicke, 1991; also see UNIT 6.4). Because aggregation is an intermolecular event, it is concentration-dependent. In this step the protein concentration is increased very slowly, allowing the protein already in solution to fold before additional unfolded protein is added. It is assumed that folded protein does not participate in aggregation.

12. Concentrate solution to ~1 mg/ml protein using a stirred cell with Diaflo PM 10 ultrafiltration membrane. Filter concentrated solution (25 to 30 ml volume) through a Millex-GV 0.22-μm filter.

    If it is not be be used immediately, the concentrated solution can be stored at −80°C for several months.

    The largest stirred cell produced by Amicon has a 2-liter maximum capacity (i.e., the Model 2000). This cell can be used to concentrate the solution to ~100 ml, and this must be further concentrated using a smaller cell (e.g., with a 200- or 400-ml capacity). For volumes >2 liters, or for a more rapid process, the Minitan tangential-flow membrane system (Millipore) can be used. This system allows a concentration rate of 0.5 to 1.0 liter/hr.

Cleave histidine tag with thrombin

13. Dilute folded protein from step 12 1:1 with 50 mM Tris·Cl (pH 7.5)/10 mM CHAPS (~0.5 mg/ml final protein concentration).

    This dilution is carried out because thrombin digestion (in step 14) works best at NaCl concentrations <200 mM.

14. Add 300 μl of 2000 NIH unit/ml thrombin (600 NIH units) to 60 ml of diluted protein from step 13 (30 mg IN^{50-212}). Incubate 30 min at 28°C with occasional mixing. Add an additional 300 μl of 2000 NIH unit/ml thrombin and incubate another 30 min. Cool solution on ice to 4°C and proceed immediately to step 15.

    Digestion conditions are optimized by conducting small-scale digestions (using <1.0 ml of protein solution from step 13). See Critical Parameters and Troubleshooting for details. The substrate/enzyme ratio used here is ~0.02% (w/w).

    Thrombin digestions are more reproducible when carried out on a relatively small scale—e.g., ~20 to 30 mg protein per batch. If more cleaved protein is required, several batches can be processed concurrently.

Remove excess thrombin by affinity chromatography

15. Pack a 1 × 10–cm column with *p*-aminobenzamidine immobilized on Sepharose 6B. Equilibrate with IN column buffer F.

16. Add 7.5 ml of 9 M urea (final concentration 1 M) to the 60-ml thrombin digest prepared in step 14.

17. Apply digest to equilibrated column from step 15 at a flow rate of 2 ml/min. Collect

BOOTH/CENTRAL UNIVERSITY LIBRARY

the column flowthrough, then wash with IN column buffer F until A_{280} is close to the baseline (this usually requires 1 to 2 column volumes).

$IN^{50\text{-}212}$ cannot be separated from thrombin by gel filtration because the $IN_{50\text{-}212}$ (mol. wt. 18,200) undegoes self-association (Hickman et al., 1994), and therefore elutes with a molecular weight close to that of thrombin (36,000). Affinity chromatography using the thrombin inhibitor p-aminobenzamidine is therefore employed.

The 1 M urea is added to the sample and column buffers to prevent nonspecific binding of $IN^{50\text{-}212}$ to the column.

The affinity column may be regenerated by washing with several column volumes of 6 M guanidine·HCl in water to remove the bound thrombin; alternatively, the thrombin can be eluted under native conditions with 50 mM Tris·Cl (pH 7.8)/10 mM benzamidine·HCl.

18. Pool flowthrough and wash fractions, add 0.1 mM AEBSF, then concentrate to ~ 2 mg/ml (in ~10 ml) using a stirred cell with Diaflo PM 10 ultrafiltration membrane.

Peform final purification and assay protein

19. Purify protein by gel filtration as in step 10, using IN column buffer G in place of IN column buffer E for equilibration and elution.

The column size can be reduced to 2.6 × 60 cm for fractionation of smaller (~30 mg) quantities of protein; the sample volume should be 5 to 10 ml.

20. Pool $IN^{50\text{-}212}$-containing fractions and filter sterilize with Millex-GV 0.22-μm filter.

The final protein solution can be concentrated if required. However, it should be noted that when CHAPS-containing solutions are concentrated by ultrafiltration, the detergent will also be concentrated. The excess detergent can be slowly removed by dialysis (APPENDIX 3B; the critical micellar concentration of CHAPS is 8 mM and its micellar size is ~8 kDa). Unfortunately, radiolabeled CHAPS is not available for monitoring the removal of that detergent; however, an enzymatic assay is available (Talalay, 1960; Coleman et al., 1979).

21. Measure A_{280} of protein solution in a cuvette with a 1-cm path length.

Native $IN^{50\text{-}212}$ (after the His tag has been cleaved) at a concentration of 1 mg/ml will have an A_{280} = 1.54.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

BGH break buffer A

100 mM Tris·Cl, pH 7.8 (pH as determined at 4°C)
20 mM EDTA (from 0.5 M stock)
10% (w/v) sucrose (100 g/liter)
200 μg/ml lysozyme (200 mg/liter)
Prepare immediately before use

BGH break buffer B

100 mM Tris·Cl, pH 7.8 (pH as determined at 4°C)
10 mM EDTA (from 0.5 M stock)
0.5 mM phenylmethylsulfonyl fluoride (PMSF)
5 mM benzamidine·HCl (780 mg/liter)
Prepare immediately before use

PMSF can be added as 2.5 ml/liter of 0.2 M PMSF stock solution.

PCU IASI CENTRAL UNIVERSITY LIBRARY

BGH column buffer A

60 mM ethanolamine·HCl, pH 9.0
5% (w/v) sucrose (50 g/liter)
Prepare immediately before use

Ethanolamine (molecular weight 61.08; density 1.02) is a strong base (pK_a = 9.4). The pH 9.6 buffer can made by dissolving 3.7 ml liquid ethanolamine free base or 5.9 g ethanolamine·HCl salt (molecular weight 97.5) per liter, and adjusting the pH to 9.6 with HCl or NaOH.

BGH column buffer B

Solution A: 0.5 M (62 g/liter) anhydrous Na_2HCO_3 (mol. wt. 84)
Solution B: 0.5 M (53 g/liter) anhydrous Na_2CO_3 (mol. wt. 106)
Mix 52 ml solution A and 148 ml solution B and dilute to 1 liter with water.

The solution may also be made directly by dissolving 7.8 g $NaHCO_3$ and 2.2 g Na_2CO_3. The 0.5 M stock solutions may be stored at least 1 month at 4°C.

BGH extraction buffer

To 420 ml H_2O add (in the order indicated):
17.2 ml 150 g/liter (2 M) glycine stock (50 mM final; store stock at 4°C)
7.8 ml 2 M NaOH
7.7 g reduced glutathione (GSH) *or* 50 ml 0.1 M GSH stock (see recipe; 5 mM final)
764 g guanidine·HCl (8 M final)
H_2O to 1 liter
Prepare immediately before use

Because of the volume increase on addition of guanidine·HCl (i.e., 1 g increases the volume 0.76 ml), buffer components must be added to 420 ml water and total volume adjusted to 1 liter at the end.

The buffer should be pH 9.6 as determined at 4°C; if it is not, it can be adjusted with 2 M NaOH or 2 M HCl. If solid glutathione (which is acidic) is added, pH will require adjustment.

If high-quality guanidine·HCl is used, the solution will be colorless and clear (see APPENDIX 3A).

BGH folding buffer A

To ~500 ml H_2O add (in the order indicated):
17.2 ml of 150 g/liter (2 M) glycine stock (50 mM glycine final; store stock at 4°C)
7.8 ml 2 M NaOH
100 g sucrose (10% w/v final)
2 ml 0.5 M EDTA (1 mM final)
310 mg reduced glutathione (GSH) *or* 10 ml of 0.1 M GSH stock (see recipe; 1 mM final)
60 mg oxidized glutathione (GSSG) *or* 1 ml of 0.1 M GSSG stock (see recipe; 0.1 mM final)
240 g urea (4 M final)
Add 4°C H_2O to 1 liter
Prepare immediately before use

The buffer should be pH 9.6 as determined at 4°C; if it is not it can be adjusted with either 2 M NaOH or 2 M HCl.

Because of the volume increase on addition of urea (i.e., 1 g increases the volume 0.76 ml), buffer components must be added to ~500 ml and total volume adjusted to 1 liter at the end. There is no need to deionize urea if ultrapure-grade material is used.

BOWMAN / CENTRAL UNIVERSITY LIBRARY

BGH folding buffer B

60 mM ethanolamine·HCl, pH 9.6 (pH as determined at 4°C)
10% (w/v) sucrose (100 g/liter)
1 mM EDTA (from 0.5 M stock)
0.1 mM reduced glutathione (GSH)
0.01 mM oxidized glutathione (GSSG)
Prepare immediately before use

GSH and GSSG can be added from stock solutions (1 ml of 0.1 M GSH and 0.1 ml of 0.1 M GSSG per liter) or as solids. If solid glutathione (which is acidic) is added, pH may require adjustment.

Ethanolamine (molecular weight 61.08; density 1.02) is a strong base (pK_a = 9.4). The pH 9.6 buffer can made by dissolving 3.7 ml liquid ethanolamine free base per liter or 5.9 g ethanolamine·HCl salt (molecular weight 97.5) per liter, and adjusting the pH to 9.6 with HCl or NaOH. It can be stored several days at 4°C.

BGH wash buffer A

100 mM Tris·Cl, pH 7.8 (pH as determined at 4°C)
2% (w/v) Triton X-100 (20 g/liter)
4 M urea (240 g/liter)
5 mM EDTA (from 0.5 M stock)
0.5 M DTT (77 mg/liter)
Prepare immediately before use

When preparing solution, allow for volume increase on addition of urea; 1 g urea will increase solution volume by 0.76 ml. There is no need to deionize urea if ultrapure-grade material is used.

Triton X-100 is a viscous liquid at 20° to 25°C; for the use intended here, there is no need for highly purified Triton.

BGH wash buffer B

100 mM Tris·Cl, pH 7.8 (pH as determined at 4°C)
0.5 mM DTT (77 mg/liter)
Prepare immediately before use

It is preferable to add DTT as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

Glutathione stock solutions, 0.1 M

Reduced glutathione: Dissolve 30.73 mg/ml reduced glutathione (GSH) in water and adjust pH to 7 with NaOH. Store at or below –20°C (stable ≥1 month).

Oxidized glutathione: Dissolve 61.3 mg/ml oxidized glutathione (GSSG) in water and adjust pH to 7 with NaOH. Store at or below –20°C (stable ≥1 month).

hIL-2 break buffer

0.1 M Tris·Cl, pH 7.5 (pH as determined at 4°C)
50 mM EDTA (from 0.5 M stock; *APPENDIX 2E*)
Prepare immediately before use

IN break buffer

50 mM Tris·Cl, pH 7.5 (pH as determined at 4°C)
5 mM EDTA (from 0.5 M stock)
5 mM benzamidine·HCl (780 mg/liter)
5 mM DTT (770 mg/liter)
0.1 mg/ml lysozyme (100 mg/liter)
Prepare immediately before use

It is preferable that DTT be added as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

BGU IASI / CENTRAL UNIVERSITY LIBRARY

IN column buffer A

10 mM Tris·Cl, pH 7.5 (pH as determined at 4°C)
4 M guanidine·HCl (382 g/liter)
1 mM 2-mercaptoethanol (2-ME; 0.070 ml/liter)

Because of the volume increase on addition of guanidine·HCl (i.e., 1 g increases the volume 0.76 ml), buffer components must be added to 420 ml water and total volume adjusted to 1 liter at the end.

Concentrated 2-ME (14.3 M) should be stored at 4°C and opened bottles replaced every 2 or 3 months.

IN column buffer B

To 500 ml H_2O add:
10 ml 1 M Tris·Cl, pH 8.0 (pH as determined at 4°C; 10 mM final)
0.73 g $NaH_2PO_4·H_2O$ (monobasic; mol. wt. 137.99)
13.44 g Na_2HPO_4 (dibasic; mol. wt. 141.96)
0.07 ml 2-mercaptoethanol (2-ME; 1 mM final)
After salts have dissolved add:
573.2 g guanidine·HCl (6 M final)
Bring temperature to 4°C and add 4°C H_2O to 1 liter
Adjust pH to 8.0 (as determined at 4°C) if necessary with 2 M NaOH or 2 M HCl
Prepare immediately before use

Because of the volume increase on addition of guanidine·HCl (i.e., 1 g increases the volume 0.76 ml), buffer components must be added to ~500 ml and total volume adjusted to 1 liter at the end.

Concentrated 2-ME (14.3 M) should be stored at 4°C and opened bottles replaced every 2 or 3 months.

IN column buffer C

Titrate IN column buffer B (see recipe) to pH 6.4 (as determined at 4°C) with HCl.
Prepare immediately before use.

IN column buffer D

Prepare 100 mM (13.8 g/liter) $NaH_2PO_4·H_2O$. If necessary adjust to pH 4.5 with 2 M NaOH or 2 M HCl.

IN column buffer E

Immediately before use, prepare IN column buffer B (see recipe), replacing the 2-ME with 5 mM DTT.

It is preferable that DTT be added as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

IN column buffer F

50 mM Tris·Cl, pH 7.5 (pH as determined at 4°C)
0.25 M NaCl (14.6 g/liter)
10 mM 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS; mol. wt. 614.9; 6.2 g/liter)
1 mM DTT (154 mg/liter)
1 mM EDTA (prepare from 0.5 M stock)
1 M urea (60 g/liter)
Prepare immediately before use

It is preferable that DTT be added as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

IN column buffer G

Prepare IN column buffer F (see recipe), but omit urea.

BOSTON CENTRAL UNIVERSITY LIBRARY

IN extraction buffer

10 mM Tris·Cl, pH 7.5 (pH as determined at 4°C)
8 M guanidine·HCl (764 g/liter)
5 mM DTT (770 mg/liter)
Prepare immediately before use

Because of the volume increase on addition of guanidine·HCl (i.e., 1 g increases the volume 0.76 ml), buffer components must be added to 420 ml and total volume adjusted to 1 liter at the end.

If high-quality guanidine·HCl is used, the solution will be colorless and clear (see APPENDIX 3A).

It is preferable that DTT be added as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

IN folding buffer

50 mM Tris·Cl. pH 7.5 (pH as determined at 4°C)
0.5 M NaCl (29.2 g/liter)
10 mM 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS; mol. wt. 614.9; 6.2 g/liter)
2 mM DTT (308 mg/liter)
Prepare immediately before use

It is preferable that DTT be added as a solid, but it may also be added as 0.2 M stock in water (30.8 g/liter; store at or below –20°C; stable ≥1 month).

RP-HPLC solvents A, B, and C

Solvent A: Prepare 0.1% (v/v) acetonitrile in water (both HPLC grade; e.g., Sigma).

Solvent B: Prepare 0.1% (v/v) TFA in acetonitrile (both HPLC grade)

Solvent C: Mix 800 ml acteonitrile and 1 ml TFA (both HPLC grade). Dilute with HPLC-grade water to 1 liter.

Prepare all solutions fresh; filter and degas before use.

Thrombin, 2000 NIH units/ml

Immediately before use reconstitute contents of a 1030 NIH-unit vial (2100 NIH units/mg) of thrombin from human plasma (Sigma) with 0.5 ml water to give a solution containing 50 mM sodium citrate, pH 6.5, 0.15 M NaCl, and 2000 NIH units/ml thrombin.

The A_{280} of 1 mg/ml thrombin is 1.83 in a cuvette with a 1-cm path length. Store the lyophized powder at or below –20°C.

COMMENTARY

Background Information

Bovine growth hormone (BGH)

There has been much commercial interest over the past fourteen years in producing large quantities of recombinant growth hormones. Human growth hormone is used as a pharmaceutical agent for treating dwarfism. Nonhuman growth hormones, including bovine and porcine growth hormones, are used in agriculture for stimulating growth and lactation in animals. These hormones were among the first group of proteins pursued by the biotechnology industry; they had been investigated decades before by protein chemists.

Bovine growth hormone (BGH; also known as bovine somatotropin) is a pituitary hormone, and the mature protein contains 191 residues (mol. wt. 21,876) with two disulfide bonds—one between residues 51 and 163 and another between residues 180 and 188. The closely related porcine growth hormone (pGH) was the first growth hormone for which the three-dimensional structure was determined (Abdel-Meguid et al., 1987). Since then, the structure of the human growth hormone (hGH) as a complex with its receptor has also been determined (deVos et al., 1992). It should be noted that hGH and pGH share 68% and 91% sequence similarity, respectively, with BGH.

FGU IASI/CENTRAL UNIVERSITY LIBRARY

In Basic Protocol 1, conditions are chosen that allow the unfolded, reduced protein to be folded and oxidized at relatively high protein concentrations (~1 to 2 mg/ml). The procedure is based on the original work of Goldberg (Orsini and Goldberg, 1978; see also references therein), who pioneered the use of cosolvents (e.g., urea in Basic Protocol 1; also see UNIT 6.4 and Maeda et al., 1994) to help maintain solubility during refolding. The use of redox buffer systems containing low-molecular-weight thiol-disulfide pairs (e.g., reduced and oxidized glutathione) to enhance the rate of disulfide-bond formation during the folding of reduced protein is based on Wetlaufer's work (Wetlaufer, 1984; see also references therein).

Although the methodology of Basic Protocol 1 was developed in 1982-1983, it remains a fairly typical example of a technique for preparative protein folding. Details of BGH expression in *E. coli* can be found in Wingfield et al. (1987a).

BGH folding and thiol oxidation

The structure of both porcine and human growth hormone consists of four antiparallel α-helices (Abdel-Meguid et al., 1987). The topology of BGH is almost certainly similar to that of pGH, due to the high degree of sequence and functional similarity. In pGH, one disulfide bond is formed between Cys-53 (found in the middle of the region connecting helices 1 and 2) and Cys-164 (located in helix 4). The other disulfide bond links the C-terminal region (at Cys-189) to helix 4 (at Cys-181).

Examination of the solution conditions for folding and oxidation of the reduced pituitary-derived BGH has shown that reduced and unfolded protein—in 6 M guanidine·HCl diluted into 4.5 M urea—can be air-oxidized with a high (90%) yield (Holzman et al., 1986). These conditions are similar to those used in Basic Protocol 1, except that in the latter, the protein concentration is >5-fold higher (>1 mg ml) and a redox buffer is used.

For further information on in vitro folding of BGH and the closely related pGH, consult Brems and Havel (1989) and Bastiras and Wallace (1992). Havel et al. (1989) present a concise and useful review of the some of the spectroscopic methods used to monitor protein conformation and structure (e.g., UV absorption, circular dichroism, and fluorescence) with examples of studies on BGH.

Comparison of authentic and recombinant BGH

Natural BGH is purified from bovine pituitaries, and examples of purifications leading to crystallizable protein are described by Bell et al. (1985) and Spitsberg (1987). The method that appears to give the most homogeneous product (Wood et al., 1989) involves extraction of pituitaries with 4.5 M urea (pH 8.8) followed by purifications employing anion- and cation-exchange chromatography in the presence of urea. The urea maintains solubility of the protein during purification; at the concentration used (4.5 M) it does not perturb or unfold the protein. The urea is finally removed from the purified protein by dialysis at pH 10. A single bovine pituitary yields ~16 mg of pure protein by this method.

Preparations of the natural protein often show varying degrees of N-terminal heterogeneity, depending on the method of purification. This may involve truncations of either one (ΔA-1) or four (A-4) residues—e.g., **(Ala)PheProAla**MetSerLeuSerGlyLeuCys AlaPhe...—where the residue in parentheses is lost in a Δ A-1 truncation and the four boldface residues are lost in a A-4 truncation. An example of a A-4 truncation is shown in Figure 6.5.1, lane E, where the two closely migrating bands at 23 and 22 kDa correspond to full-length and N-terminal-deleted (A-4) protein, respectively (for further details see Wingfield et al., 1987b, and references therein; also see Langley et al., 1987a). The two forms can be resolved under native conditions by chromatofocusing, which takes advantage of the small difference in their isoelectric points; full-length and A-4 BGH have pI values of 8.2 and 8.0, respectively (Wingfield et al., 1987b). Protein purified by the method of Wood et al. (1989) has a single and correct N-terminus.

Recombinant hormones have been expressed that contain the full BGH sequence (e.g., Langley et al., 1987a), as well as others with the ΔA-1 deletion (e.g., Wingfield et al., 1987a; Bogosian et al., 1989) and the A-9 deletion (Wingfield et al., 1987a). Purified A-1 BGH retains the initiating Met at the N-terminus (MetPhePro...), but when the initiating Met is placed in front of the Ala, the Met is removed to give the authentic sequence (AlaPhePro...). Deletion of at least the first nine N-terminal residues does not appear to effect the biological activity of BGH.

Another source of chemical heterogeneity in both authentic and recombinant growth hor-

BOSTON UNIVERSITY LIBRARY

mones is the deamidation of specific Asn and Gln residues (Secchi et al., 1986). This is partly responsible for the multiple band patterns observed in isoelectric focusing gels (Wingfield et al., 1987a; 1987a). In typical authentic and recombinant BGH preparations, ~20% of the protein is deamidated. However, because these modified forms have lower pIs than unmodified protein, they can be separated by chromatofocusing (Wingfield et al., 1987a). The cation-exchange chromatography step (at pH 6.9 in 4.5 M urea) used by Wood et al. (1989) for purification of pituitary BGH also appears to separate deamidated from unmodified protein. Deamidated protein is more positively charged than unmodified protein, and will therefore bind more tightly to a negatively charged cation-exchange matrix.

Authentic and recombinant BGH both exhibit self-association with the characteristics of a rapid monomer-dimer equilibrium. The dissociation constant (K_d) for authentic BGH is 6.6 × 10–6 M (Fernandez and Delfino, 1983), which is similar to that determined for both the A-1 and A-9 recombinant forms of BGH (P.T.W., unpub. observ.). At concentrations of 1.0 and 0.1 mg/ml, ~76% and 42% of the protein will be dimeric, respectively. Under physiological conditions, the monomer appears to be the active species. Monomeric protein binds to two molecules of the extracellular domain of the cell-surface receptor, forming a 1:2 molar complex analogous to the human growth hormone (Staten et al., 1993).

Other approaches used to extract and fold recombinant BGH

Langley et al. (1987a) extracted protein from inclusion bodies using 6 M guanidine·HCl (pH 8.0) as the denaturant. The extract was allowed to oxidize in air >72 hr at room temperature in the presence of the denaturant. Oxidized protein was fractionated by gel filtration in guanidine·HCl, the fractions making up the monomer peak pooled, and the protein folded by removal of denaturant using dialysis. The rationale for this method is that, in a protein with four cysteine residues that form two disulfide bonds (such as BGH), there are three possible intramolecular disulfide linkage combinations. Hence, even if the oxidation proceeds in a random manner as expected for unfolded protein, at least 33% of the protein produced will have the correct pattern. In practice, the recovery of correctly folded oxidized protein was significantly higher than 33%—suggesting that even in 6 M guanidine·HCl, BGH may contain some elements of native-like structure. This method is geared toward large-scale protein production and gives good protein recovery. However, the long incubation period in guanidine·HCl appears to result in ~50% of the protein being deamidated (Langley et al., 1987b).

Bogosian et al. (1989) used 4.5 M urea at pH 10.7 to extract BGH from inclusion bodies. The extract was then stirred for 48 hr, allowing the protein to oxidize in air. This method is mentioned because it results in formation of a novel concatenated dimer, formed by the interlocking of disulfide loops (Violand et al., 1989). It is most likely that under the conditions used for extraction, the protein folding and oxidation is initiated from partially folded associated protein.

Ovine growth hormone (which is closely related to BGH) was folded and purified with an overall recovery of 30% by a procedure similar in approach to Basic Protocol 1 (Wallis and Wallis, 1989). This method, however, used 2 M guanidine·HCl instead of urea as the co-solvent to maintain solubility during folding. It also employed air oxidation for 24 hr instead of the redox buffer system.

Interleukin 2

Interleukin 2 (IL-2) is a member of the cytokine family. It is composed of a compact core bundle of four antiparallel α-helices (Bazan, 1992; McKay, 1992). IL-2 plays important roles in the proliferation and differentiation of T lymphocytes as well as in regulation of the immune system. The biological effects of IL-2 are mediated via binding to IL-2 receptors (Minami et al., 1993).

The human IL-2 (hIL-2) gene has been cloned and expressed in *E. coli* as insoluble inclusion bodies (Devos et al., 1983). The purification of recombinant hIL-2 has been described and the protein thoroughly characterized (Liang et al., 1985; Kato et al., 1985; Weir and Sparks, 1987). Basic Protocol 2 details the solubilization, refolding, and purification of recombinant hIL-2. This procedure can be scaled up to yield hIL-2 on the scale of hundreds of milligrams.

hIL-2 has three cysteines; hence, three possible disulfide-linked forms can exist. The native isomer consists primarily of one form—i.e., the one that contains a disulfide linkage between Cys-58 and Cys-105 (Robb et al., 1984; Wang et al., 1984; Liang et al., 1985).

PCU IASWER TRAI UNIVERSITY LIBRARY

Reversed-phase-HPLC has been used to separate these isomers (Browning et al., 1987; also see Fig. 6.5.3).

HIV-1 integrase

An early part of the life cycle of all retroviruses (including the human immunodeficiency virus, HIV) is the integration of a DNA copy of the viral genome into the host chromosome. This step is essential for viral replication. Retroviral DNA integration is carried out by a defined set of DNA cutting and joining reactions, catalyzed by integrase protein (Katz and Skalka, 1994).

The HIV-1 integrase is a 288-residue protein (mol. wt. 32,200) which has been expressed in *E. coli* (Sherman and Fyfe, 1990). The protein is located in the pellet obtained by low-speed centrifugation following cell breakage, but is apparently not highly aggregated into inclusion bodies, as it can be extracted into high-ionic-strength solution (e.g., 1 M NaCl) without a denaturant. The insolubility of the protein is probably a result of nonspecific binding to *E. coli* nucleic acid. The salt-extraction procedure was originally used by Terry et al. (1988) for extraction of avian sarcoma-leukosis virus integrase expressed in *E. coli*.

Dissection of the HIV-1 integrase by preparation of a series of deletion mutants (Bushman et al., 1993; see also references therein) demonstrated that a central core region of residues 50 to 212 was enzymatically active, carrying out a subset of the reactions catalyzed by the full-length enzyme. This deletion mutant is interesting for structural studies as it has better solubility than the full-length enzyme, which is notoriously difficult to handle as a result of limited solubility in the usual aqueous solvents.

The HIV-1 integrase deletion mutant ($IN^{50\text{-}212}$) is expressed in *E. coli* as a fusion protein with the N-terminal extension sequence GlySerSerGlyHisHisHisHisHisHisSerSer GlyLeuVal**ProArgGly**SerHisMet. This sequence (a His tag) contains a six-residue histidine repeat which is responsible for the selective high-affinity binding of the fusion protein to a nickel-chelate column. The boldface portion of the His tag sequence indicates the location of the specific thrombin-cleavage site between Arg-16 and Gly-17, which is exploited in removal of the tag using thrombin.

In Basic Protocol 3 the His-tagged $IN^{50\text{-}212}$ is expressed as an insoluble (i.e., inclusion-body-type) protein, and is purified under denaturing conditions, taking advantage of the fact that metal-chelate affinity chromatography (MCAC) can be carried out in the presence of protein denaturants such as guanidine·HCl. Following purification, the protein is folded and the His tag removed by exploiting the specific thrombin-cleavage site of the tag sequence (underlined above). The cleaved protein has the N-terminal sequence GlySerHisMetHisGlyGln.... The Met at position 4 corresponds to Met 50 of the wild-type integrase and residues 1 to 3 (GlySerHis) are derived from the tag. Excess thrombin is removed by affinity chromatography using an immobilized inhibitor (*p*-aminobenzamidine). In the last step of the purification, the $IN^{50\text{-}212}$ is subjected to gel-filtration to remove any aggregated multimers. The protein produced is 166 residues long, the molecular weight estimated from the DNA coding sequence is 18,200, and the calculated isoelectric point is ~7.

Although the folded deletion mutant has enhanced solubility as compared with the wild-type enzyme, the zwitterionic detergent CHAPS is included in buffers to maintain solubility, especially during protein folding. Following protein purification the detergent can be removed or exchanged for other buffer additives.

Substitution of Phe-185 for a Lys residue appears to increase the solubility of $IN^{50\text{-}212}$. This mutant protein, expressed in *E. coli* with the N-terminal His tag, was extracted with 1 M salt and purified by MCAC. The extraction with high-ionic-strength salt solution in the absence of denaturant is analogous to the procedure used with the full-length protein, discussed above. The crystal structure of this protein has been determined at 2.5-Å resolution. The overall topology consists of five β-sheets flanked by helical regions, showing that the integrase domain belongs to a superfamily of polynucleotidyl transferases that includes ribonuclease H (Dyda et al., 1994).

Critical Parameters

Bovine growth hormone (BGH)

Protein extraction. The protein should be extracted from the inclusion bodies in a monomeric and fully reduced state. This provides a defined starting point from which to develop a reproducible folding protocol. Based on pilot-scale experiments in which washed pellets (from step 4 of Basic Protocol 1) were extracted with various protein denaturants (P.T.W., unpub. observ.; also see *UNIT 6.3*), 8 M guanidine·HCl was chosen for solubilization of aggregated BGH. The most effective ones were

PRINCETON UNIVERSITY LIBRARY

urea and guanidine·HCl at 6 M to 8 M concentrations, each of which solubilized >80% of the protein. The extracts were further analyzed by gel filtration on Sephacryl S-300 in the presence of the same concentration of the particular denaturant used for the initial extraction. The results indicated that guanidine·HCl at concentrations >6 M yielded solubilized protein of which >80% was monomeric. Although 8 M urea also solubilized >80% of the protein, only ~20% of this protein was monomeric, the remainder consisting of dimers and higher-order aggregates.

Protein folding and oxidation. Basic Protocol 1 describes an empirical approach based on some of the principles discussed in *UNITS 6.1 & 6.4*. The determination of basic physicochemical/conformational properties for a particular protein will take some of the guesswork out of developing a suitable folding protocol. For BGH, purified authentic protein from bovine pituitary glands was available, and had been fairly well characterized before recombinant proteins were produced. This situation is not common, as the natural counterparts of many recombinant proteins are rare and may never have been purified. The optimal ratio of reduced to oxidized glutathione in the folding buffers used in Basic Protocol 1 was empirically determined. However, it is commonly observed that the highest folding yields are obtained when the ratio of GSH/GSSG is between 5 and 10 and the total reduced and oxidized glutathione concentration is in the range of 1 to 10 mM.

Deamidation. As mentioned above, BGH is susceptible to deamidation. Although deamidated forms (charge isomers) can be resolved by ion-exchange methods (Wingfield et al., 1987a), their formation should be minimized. This is achieved by limiting the amount of time the protein spends in the fully unfolded state, especially at pH 9.5. For example, the guanidine·HCl extract should be processed immediately and diluted as described in step 6 of Basic Protocol 1. The dialysis steps (at pH 9.5) should be performed for the times indicated, not longer. The folded protein is stable and does not appear to deamidate spontaneously on storage.

Scaleup. For larger-scale production, a Manton-Gaulin-APV homogenizer (*UNIT 6.1*) is used to break cells instead of the French press. In addition, a larger Sephadex G-100 column (e.g., 5.0 × 90–cm) is used in step 12 of Basic Protocol 1.

Interleukin-2

Protein purification. Recombinant hIL-2 should be extracted as a monomer (see Basic Protocol 2, step 5). The extraction has been tested with various concentrations of acetic to determine the optimal condition; 20% acetic acid has been found to extract the maximum amount of monomeric hIL-2 from inclusion bodies. However, in the gel-filtration step (see Basic Protocol 2, step 8), 20% acetic acid is too corrosive for the metal parts of the chromatography system (e.g., the stands, pump, and fraction collector); thus 10% acetic acid is recommended.

Protein folding. The hIL-2 monomer that elutes from the Sephadex G-100 column (see Basic Protocol 2, step 8) yields two peaks upon analytical reversed-phased chromatography (see Support Protocol). These two peaks represent the refolded, oxidized hIL-2 and unfolded, reduced hIL-2. The proportion of the oxidized form is increased by the dialysis step (see Basic Protocol 2, step 9). The time and volume recommended in the dialysis step have been determined empirically to produce the best yield of oxidized hIL-2. Longer periods of dialysis (e.g. 48 hr) causes do not increase the yield of oxidized hIL-2.

Scaleup. The parameters for larger-scale production are the same as those for BGH (see above).

HIV-1 integrase

MCAC. It is important to use low concentrations of 2-mercaptoethanol in the sample and column buffer (1 mM is safe) and to completely avoid dithiothreitol, as either reductant will strip Ni^{2+} from the MCAC matrix (see *UNIT 9.4*).

Protein folding. Because HIV-1 integrase is not a very soluble protein and has a high tendency toward self-association and aggregation (Hickman et al., 1994), the solubility and stability of the folded protein must be maintained by including a relatively high salt concentration as well as the detergent CHAPS in the folding buffer and in all subsequent column buffers. The buffer additives indicated for HIV-1 integrase are usually not required for folding of the average protein.

Thrombin cleavage. It is necessary to perform pilot-scale experiments to optimize the conditions for removing the His tag using thrombin. The usual parameters to vary are enzyme/substrate ratios and incubation time and temperature. Thrombin is a serine protease, and can be irreversibly inhibited with either PMSF or AEBSF. Removal of the tag can be

FGU IASI CENTRAL UNIVERSITY LIBRARY

monitored by SDS-PAGE (*UNIT 10.1*), in which case the thrombin should be quenched with 1 mM PMSF or AEBSF prior to addition of SDS and heating. There is an ~2-kDa mass difference between tagged and nontagged protein (see Fig. 6.5.2). The digestion should be as complete as possible at the specific cleavage site (see Background Information); overdigestion may result in nonspecific cleavage at other sites. It is also important to avoid denaturation of the (thrombin) (e.g., by vigorous stirring or overheating), as partially denatured proteases often exhibit some loss of specificity, cleaving at unpredictable sites.

Troubleshooting

Bovine growth hormone (BGH)

Most precautions and guidelines for folding and oxidation are given in annotations to the individual steps of Basic Protocol 1. It should be cautioned that, when using a mixture of oxidized and reduced glutathione for oxidizing proteins that contain both free and disulfide-linked cysteines in the native state (e.g., IL-2), the potential exists for glutathionylation, in which the unpaired cysteine(s) form a mixed disulfide with glutathione. This modification will cause charge heterogeneity, and can be readily detected by electrospray ionization mass spectrometry (ESI-MS), which will indicate a mass increase of 305.3 for each GSH moiety incorporated.

Interleukin-2

Low-molecular-weight impurities extracted together with monomeric hIL-2 in acetic acid (see Basic Protocol 2, step 5) are usually eluted earlier than the oxidized form of hIL-2 in RP-HPLC (see Basic Protocol, step 12) and can thus be removed. If the RP-HPLC peak containing the oxidized form of hIL-2 is contaminated by low-molecular-weight impurities, SDS-PAGE analysis of each fraction composing the peak is recommended to avoid pooling fractions containing impurities. Additional information regarding troubleshooting is given in annotations to individual steps of Basic Protocol 2.

HIV-1 integrase

MCAC. Protein will not bind to the column if the His tag has been nonspecifically degraded or removed by *E. coli* proteases. SDS-PAGE (*UNIT 10.1*) of the cell extracts and fractions during the purification should indicate if the tag is present. The His tag adds ~2 kDa to the apparent mass of the protein. The denaturant concentration should be high enough to maintain solubility of the protein during chromatography. Insoluble or aggregated protein will not bind to the column and will be located in the column flowthrough or will clog the top of the matrix. If protein aggregates on the column, a higher concentration of denaturant should be tried, although aggregation should not happen using 6 M guanidine·HCl in column buffer A. Further details on troubleshooting the use of the Ni-NTA resin are given in the manufacturer's literature (Qiagen, 1992) and in *UNIT 9.4*.

Thrombin cleavage. If undigested (His-tagged) protein still remains after digestion, the standard procedure is to reapply the mixture to an MCAC column, whereupon the matrix will specifically bind the unprocessed protein containing the His tag. The processed (cleaved) protein is not bound and elutes in the column flowthrough. This approach, although satisfactory for monomeric proteins, is complicated with multimeric proteins and those that exhibit reversible self-association—e.g., HIV-1 integrase. For example, in a partially digested dimeric protein, the subunits can be arranged in three possible configurations—TT, TM, and MM—where T is still tagged and M has had the tag cleaved. It may be possible to separate TT and TM by gradient elution from the MCAC column, as the former might be expected to bind more tightly than the latter. A simpler approach is to use solvent conditions that dissociate the protein subunits without denaturing them. For HIV-1-integrase, this can be achieved with 2 to 3 M urea.

Anticipated Results

BGH

The purification of Δ A-9 BGH is summarized in Table 6.5.1. About 90 to 100 mg of protein are obtained from 50 g (wet weight) of cells with an overall yield in the range 15% to 25%. For larger and/or multidomain proteins, much lower yields (1% to 5%) may be more typical. Gel analysis of the BGH in cell extracts and in purified protein is shown in Figure 6.5.1.

Interleukin-2

If the expression of hIL-2 in *E. coli* is ~10%, 20 g of cell pellet will yield 4 to 5 mg of purified protein, with a specific activity of 5×10^6 U/mg. The purity of the hIL-2 should be >98% as indicated by analytical C18 column chromatography (see Support Protocol).

JONAS SALK INSTITUTE LIBRARY

Table 6.5.1 Purification of a Recombinant Bovine Growth Hormone Analog[a]

Stage of Purification[b]	Total Protein[c] (mg)	Specific BGH content[d] (%)	Total BGH (mg)	Yield (%)
Cells	5000	7.5	375	100
Washed pellet (step 4)	375	60	225	60
Dialysis supernatant (step 8)	200	75	150	40
DEAE-Sepharose pool (step 11)	131	95	125	33
Sephadex G-100 pool (step 14)	95	99	94	25

[a]The summary refers specifically to a biologically active analog of BGH (δ-9 BGH) in which the full-length sequence is truncated at the *N*-terminus by eight residues and serine is substituted for glycine at the first position. Similar relative yields were obtained with a δ-1 analog in which the *N*-terminal Ala of the native sequence was replaced by Met, but the amount of this protein expressed in *E. coli* was several-fold lower than that obtained for δ-9 BGH (Wingfield et al., 1987a)

[b]The numbers in parentheses refer to Basic Protocol 1 steps.

[c]Determinations were made using the Bio-Rad Protein Assay Kit, except those for the DEAE and Sephadex pools, which were made by UV absorbance measurements.

[d]Percentage of total protein that is BGH. Estimates made by densitometric scanning of Coomassie blue stained SDS-polyacrylamide gels.

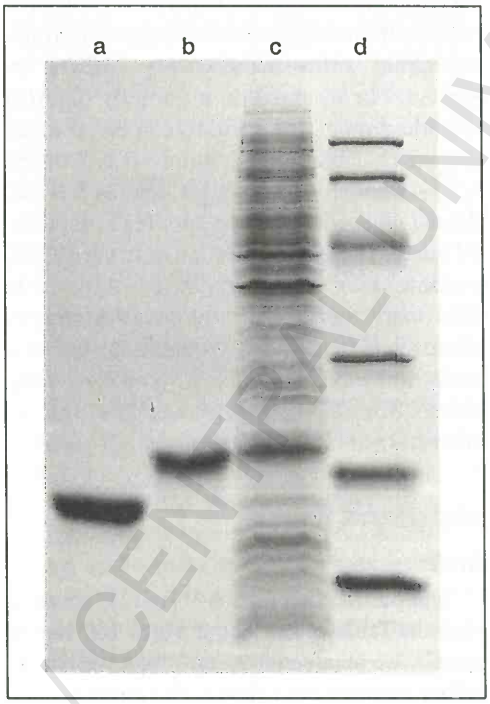


Figure 6.5.4 Results of SDS-PAGE of H1V-1 integrase$^{50\text{-}212}$ on a 12.5% polyacrylamide gel stained with Coomassie blue. Lane A, recombinant HIV-1 integrase$^{50\text{-}212}$; lane B, recombinant HIV-1 integrase$^{50\text{-}212}$ with N-terminal His tag; lane C, extract of HIV-1 integrase–expressing *E. coli* cells used for purification.

In a scaled-up process (see Critical Parameters), 25 to 30 mg of purified hIL-2 may be obtained from 200 g of *E. coli* cell pellet.

HIV-1 integrase

Figure 6.5.4 shows results of SDS-PAGE of the purified IN$^{50\text{-}212}$ before (lane B) and after (lane A) removal of the His tag by thrombin digestion. The molecular weight difference of ~2000 kDa corresponds to the removal of the 16 amino acid residues comprising the N-terminal His tag.

The overall recovery of purified protein is ~30%. For example, from 1 g (wet weight) of cells (containing ~5% IN$^{50\text{-}212}$) ~1.5 mg of purified protein are obtained (see Figure 6.5.4, lane C, for SDS-PAGE analysis of whole IN$^{50\text{-}212}$–expressing *E. coli*). Most of the steps, including protein folding, have high recoveries (80% to 95%). The yield of thrombin-digested

BGU IASI CENTRAL UNIVERSITY LIBRARY

protein from purified tagged protein is between 60% to 70%.

The enzymatic and biophysical properties of the protein are detailed in Hickman et al. (1994).

Time Considerations

BGH

Basic Protocol 1 for the purification of BGH takes 5 days. Protein can be stored at –80°C at the end of step 4 (as washed pellets) or step 10 (as pooled fractions from ion-exchange).

Day 1: Cell breakage, preparation of washed pellets and extraction with guanidine·HCl (steps 1 to 5) requires ½ day of work. The protein extract is then dialyzed overnight (step 5).

Day 2: After changing the dialysis buffer (step 7), dialysis is continued at least 6 hr. At this step the dialysis can be left overnight or directly processed by ion-exchange chromatography (steps 8 to 10), which takes 6 to 8 hr and can be run overnight.

Day 3: Ion-exchange chromatography is run, or if ion-exchange was performed on day 2, the pooled fractions are concentrated for several hours (step 11), then applied to the gel-filtration column and chromatographed for several hours (steps 12 and 13). The gel-filtration column is run is overnight if a low-pressure matrix (e.g., Sephadex G-100) is used or on the same day if a medium-pressure matrix (e.g., Superdex 200) is used.

Day 4: The protein is analyzed (e.g., by SDS-PAGE or isoelectric focusing), concentrated if required (step 14), and frozen or prepared for lypohilization.

Interleukin-2

In Basic Protocol 2, preparation and lysis of the cells (steps 1 and 2) will take 1 hr, cell breakage and preparation of washed pellets (steps 3 and 4) will take 3 hr, acid extraction (steps 5 and 6) will take 1 hr, Sephadex G-100 column chromatography (step 7) will take 12 hr, dialysis of the pooled fractions will be carried out overnight, and RP-HPLC (steps 9 to 12) will require 7.5 hr. The final dialysis is run overnight. In the Support Protocol for resolution of native and misfolded forms of hIL-2, it will take ~1 hr to run a sample after the blank runs to establish the baseline profile have been completed.

HIV-1 integrase

Basic Protocol 3 is usually carried in two stages. In stage 1, purification of the unfolded protein that still contains the His tag (steps 1 to 8) is usually carried out on a relatively large scale (using 100 g of cells). If the Superdex 200 column used in step 5 and the MCAC column used in step 6 are run using the Pharmacia Biotech Biopilot system, this stage will take 3 to 4 days to complete. It will take longer if low-pressure columns are used. In stage 2, protein folding (steps 11 and 12), removal of the His tag by thrombin cleavage (steps 13 and 14), affinity chromatography (steps 15 to 17), and gel filtration (step 19) are performed. This stage is carried out repeatedly with relatively small amounts of protein (i.e., 30 mg; this represents <10% of the protein produced in the Stage 1). Using the Pharmacia Biotech FPLC system for the chromatography steps, stage 2 takes 2 days.

Literature Cited

Abdel-Meguid, S.S., Shieh, H.-S., Smith, W. W., Dayringer, H.E., Violand, B.N., and Bentle, L.A. 1987. Three-dimensional structure of a genetically engineered variant of porcine growth hormone. *Proc. Natl. Sci. U.S.A.* 84:6434-6437.

Ackers, G.K. 1970. Analytical gel chromatography of proteins. *Adv. Protein Chem.* 24:343-446.

Bastiras, S. and Wallace, J. C. 1992. Equilibrium denaturation of recombinant porcine growth hormone. *Biochemistry* 31:9304-9309.

Bazan, J.F. 1992. Unraveling the structure of IL-2. *Science* 257:410-413.

Bell, J.A., Moffat, K., Voderhaar, B.K., and Golde, D.W. 1985. Crystallization and preliminary X-ray characterization of bovine growth hormone. *J. Biol. Chem.* 260:8520-8525.

Bogosian, G., Violand, B.N., Dorward-King, E.J., Workman, W.E., Jung, P.E., and Kane, J.F. 1989. Biosynthesis and incorporation into protein of norleucine by *Escherichia coli. J. Biol. Chem.* 264:531-539.

Bottomly, K., Davis, C.S., and Lipsky, P.E. 1991. Measurement of human and murine interleukin-2 and interleukin-4. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Marguiles, E.M. Shevach, and W. Strober, eds.) pp. 6.3.1-6.3.12. John Wiley & Sons, New York.

Brems, D.N. and Havel, H.A. 1989. Folding of bovine growth hormone is consistent with the molten globule hypothesis. *Proteins Struct. Funct. Genet.* 5:93-95.

Browning, J.L., Mattaliano, R.J., Chow, E.P., Liang, S.-M., Allet, B., Rosa, J., and Smart, J.E. 1986. Disulfide scrambling of interleukin-2: HPLC resolution of the three possible isomers. *Anal. Biochem.* 155:123-128.

BOULDER CENTRAL UNIVERSITY LIBRARY

Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P., and Craigie, R. 1993. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for the polynucleotidyl transfer and zinc binding. *Proc. Natl. Acad. Sci. U.S.A.* 90:3428-3432.

Coleman, R., Iqbal, S., Godfrey, P.P., and Billington, D. 1979. Membranes and bile formation. Composition of several mammalian biles and their membrane-damaging properties. *Biochem. J.* 178:201-208

deVos, A.M., Uttsch, M., and Kossiakoff, A.A. 1992. Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science* 255:306-312.

Devos, R., Plaetinc, G., Cheroutr, H., Simons, G., Degrave, W., Tavernie, J., Remaut, E., and Fiers, W. 1983. Molecular cloning of human interleukin-2 carrier DNA and its expression in *Escherichia coli*. *Nucl. Acids Res.* 11:4307-4323.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., and Davies, D.R. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: Similarity to other polynucleotidyl transferases. *Science* 266:1981-1986.

Edelhoch, H. and Burger, H.G. 1966. The properties of bovine growth hormone. II. Effects of urea. *J. Biol Chem.* 241:458-463.

Fernandez, H.N. and Delfino, J.M. Covalent crosslinking of the bovine somatotropin dimer. *Biochem. J.* 209:107-115.

Goldberg, M.E., Ruldolph, R., and Jaenicke, R. 1991. A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. *Biochemistry* 30:2790-2797.

Goldenberg, D.P. 1989. Analysis of protein conformation by gel electrophoresis. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 225-250. IRL Press, Oxford.

Graf, L., Li, C.H., and Bewley, T. 1975. Selective reduction and alkylation of the COOH-terminal disulfide bridge in bovine growth hormone. *Int. J. Pep. Prot. Res.* 7:467-473.

Havel, H.A., Chao, R.S., Haskell, R.J., and Thamann, T.J. 1989. Investigations of protein structure with optical spectroscopy: Bovine growth hormone. *Anal. Chem.* 61:642-650.

Hickman, A., Palmer, I., Engelman, A., Craigie, R., and Wingfield, P. 1994. Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. *J. Biol. Chem.* 269:29279-29287.

Holzman, T.F., Brems, D.N., and Dougherty, J.J. 1986. Reoxidation of reduced bovine growth hormone from a stable secondary structure. *Biochemistry* 25:6707-6917.

Jaenicke, R. 1991. Protein folding: Local structures, domains, subunits, and assemblies. *Biochemistry* 30:3147-3161.

Kato, K., Yamada, T., Kawahara, K., Onda, H., Asano, T., Sugino, H., and Kakinuma, A., 1985. Purification and characterization of recombinant human interleukin-2 produced in Escherichia coli. *Biochem. Biophys. Res. Commun.* 130:692-699.

Katz, R.A. and Skalka, A.M. 1994. The retroviral enzymes. *Annu. Rev. Biochem.* 63:133-173.

Langley, K.E., Berg, T.F., Strickland, T.W., Fenton, D.M., Boone, T.C., and Wypych, J. 1987a. Recombinant-DNA-derived bovine growth hormone from *E. coli*. I. *Eur. J. Biochem.* 163:313-321.

Langley, K.E., Lai, P.-H., Wypych, J., Everett, R.R., Berg, T.F., Krabill, L.F., and Souza, L.M. 1987b. Recombinant-DNA-derived bovine growth hormone from *E. coli*. II. *Eur. J. Biochem.* 323-330.

Liang, S.-M., Allet, B., Rose, K., Hirschi, M., Liang, C.-M., and Thatcher, D.R. 1985. Characterization of human interleukin 2 derived from *Escherichia coli*. *Biochem. J.* 229:429-439.

Liang, S.-M., Thatcher, D., Liang, C.-M., and Allet, B. 1986. Studies of structure activity relationships of human interleukin-2. *J. Biol. Chem.* 261:334-337.

Maeda, Y., Ueda, T., Yamada, H., and Imoto, T. 1994. The role of net charge on the renaturation of reduced lysozyme by sulfhydryl-disulfide interchange reaction. *Prot. Engineering* 7:1249-1254.

McKay, D.B. 1992. Unraveling the structure of IL-2-response. *Science* 257:412-413.

Minami, Y., Kono, T., Miyazaki, T., and Taniguchi, T. 1993. The IL-2 receptor complex: Its structure, function, and target genes. *Annu. Rev. Immunol.* 11:245-267.

Orsini, G. and Goldberg, M.E. 1978. The renaturation of reduced chymotrypsin A in guanidine HCl. *J. Biol. Chem.* 253:3453-3458.

Qiagen. 1992. The QIAexpressionist, 2nd ed. Qiagen, Chatsworth, Calif.

Riddles, P. W., Blakeley, R.L., and Zerner, B. 1979. Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid)—a reexamination. *Anal Biochem.* 94:75-81.

Robb, R.J., Kutney, R.M., Panico, M., Morris, H.R., and Chowdhry, V. 1984. Amino acid sequence and post-translational modification of human interleukin-2. *Proc. Natl. Acad. Sci. U.S.A.* 81:6486-6490.

Secchi, C., Biondi, P. A., Negri, A., Borrini, R., and Ronchi, S. 1986. Detection of desamido forms of purified bovine growth hormone. *Int. J. Pept. Prot. Res.* 28:298-306.

Sherman, P.A. and Fyfe, J.A. 1990. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleavage activity. *Proc. Natl. Acad. Sci. U.S.A.* 87:5119-5123.

Spitsberg, V.L. 1987. A selective extraction of growth hormone from bovine pituitary gland and its further purification and crystallization. *Anal. Biochem.* 160:489-495.

Staten, N.R., Byatt, J.C., and Krivi, G.G. 1993. Ligand-specific dimerization of the extracellular

ICU IASI CENTRAL UNIVERSITY LIBRARY

domain of the bovine growth hormone receptor. *J. Biol. Chem.* 268:18467-18473.

Stoscheck, C.M. 1990. Quantitation of protein. *Methods Enzymol.* 182:50-67.

Talalay, P. 1960. Enzymic analysis of steroid hormones. *Biochem. Anal.* 8:119-142.

Terry, R., Soltis, D.A., Katzman, M., Cobrinik, D., Leis, J., and Skalka, A.M. 1988. Properties of avian sarcoma-leukosis virus pp32-related polendonucleases produced in *Escherichia coli. J. Virol.* 62:2358-2365.

Thannhauser, T.W., Konishi, Y., and Scheraga, H.A. 1984. Sensitive quanititative analysis of disulfides in polypeptides and proteins. *Anal. Biochem.* 138:181-188.

Violand, B.N., Takano, M., Curran, D.F., and Bentle, L.A. 1989. A novel concatenated dimer of recombinant bovine somatotropin. *J. Prot. Chem.* 8:619-628.

Wallis, O.C. and Wallis, M. 1989. Purification and properties of recombinant DNA–derived ovine growth hormone analogue (oGH1) expressed in *Escherichia coli. J. Mol. Endocrinology* 4:61-69.

Wang, A., Lu, S.-D., and Mark, D. 1984. Studies of structure-activity relationships of human interleukin-2. *Science* 224:1431-1433.

Weir, M.P. and Sparks, J., 1987. Purification and renaturation of recombinant human interleukin-2. *Biochem. J.* 245:85-91.

Wetlaufer, D.B. 1984. Nonenzymatic formation and isomerization of protein disulfides. *Methods Enzymol.* 107:301-304.

Wingfield, P.T., Graber, P., Buell, G., Rose, K., Simona, M., and Burleigh, B.D. 1987a. Preparation and characterization of bovine growth hormones produced in recombinant *Escherichia coli. Biochem. J.* 243:929-839.

Wingfield, P.T., Graber, P., Rose, K., Simona, M.G., and Hughes, G.J. 1987b. Chromatofocusing of N-terminally processed forms of proteins: Isolation and characterization of two forms of interleukin-1β and bovine growth hormone. *J. Chromatogr.* 387:291-300.

Wood, D.C., Salsgiver, W.J., Kasser, T.R., Lange, G.W., Rowold, E., Violand, B.N., Johnson, A., Leimgruber, R.M., Parr, G.R., Siegel, N.R., Kimack, N.M., Smith, C.E., Zobel, J.F., Ganguli, S.M., Garbow, J.R., Bild, G., and Krivi, G.G. 1989. Purification and characterization of pituitary bovine somatotropin. *J. Biol. Chem.* 264:14741-14747.

Contributed by Paul T. Wingfield (BGH and HIV-1 integrase) and Ira Palmer (HIV-1 integrase)
National Institutes of Health
Bethesda, Maryland

Shu-Mei Liang (interleukin-2)
North American Vaccine Corp.
Beltsville, Maryland

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Expression and Purification of GST Fusion Proteins

This unit describes how pGEX vectors (available from Pharmacia Biotech) can be used for high-level inducible intracellular expression of polypeptides as fusions with glutathione-*S*-transferase (GST) in *Escherichia coli*. GST fusion proteins are easily purified under nondenaturing conditions by affinity chromatography (Chapter 9) using a glutathione–Sepharose 4B conjugate. The amino-terminal GST moiety can then be cleaved from the protein of interest using a specific protease cleavage site located between the GST moiety and the recombinant polypeptide. Finally, the GST moiety can be removed by rechromatographing the sample on the glutathione-Sepharose column.

Potential applications of the system include the expression and purification of large quantities of individual polypeptides for use in structural determinations using either nuclear magnetic resonance (NMR) or crystallography, immunological studies, vaccine production, and structure-function studies involving protein-protein and DNA-protein interactions.

The basic protocols are chromatography-based adaptations of the manufacturer's recommendations and include batch and column purification methods. The experience of this laboratory is that, although batch purifications may be slightly faster, column-based purification steps usually provide higher protein yields and higher purity of the final product.

GST fusion proteins can be expressed at high levels in *E. coli* grown using an environmental shaker (see Basic Protocol 1). The soluble fusion protein is purified using a single-step glutathione–Sepharose 4B affinity column (see Basic Protocol 2). Fusion proteins that are found in inclusion bodies can be extracted using denaturants and then refolded before affinity chromatography using a glutathione–Sepharose 4B column (see Alternate Protocol 1). Soluble fusion proteins can also be purified batchwise using glutathione–Sepharose 4B affinity matrix (see Alternate Protocol 2). The fusion protein can be cleaved with either thrombin or factor Xa in solution to separate the protein of interest from glutathione (see Basic Protocol 3). The protein of interest can also be released by protease digestion of the fusion protein immobilized on a glutathione–Sepharose 4B column (see Alternate Protocol 3), or by batchwise protease cleavage and separation from the resin (see Alternate Protocol 4). Affinity purification using glutathione–Sepharose 4B can be used to remove the GST moiety after enzymatic cleavage (see Support Protocol 1). HPLC gel filtration is used as a final purification step to obtain >98% pure protein (see Support Protocol 2). The relationships between the protocols are shown in Figure 6.6.1.

STRATEGIC PLANNING

A variety of pGEX expression vectors are commercially available (Pharmacia Biotech) that contain a *tac* promoter for chemically inducible, high-level protein expression. The available pGEX vectors have an open reading frame encoding glutathione-*S*-transferase (GST) followed by multiple cloning sites. These are followed by termination codons in each reading frame (Figs. 6.6.2 and 6.6.3). A fragment of DNA containing the genetic sequence for the polypeptide of interest is ligated into an appropriate pGEX vector and transformed into *E. coli*. It should be noted that although expression in *E. coli* is efficient, there is no post-translational modification machinery. Successful expression of GST

Contributed by Sandra Harper and David W. Speicher
Current Protocols in Protein Science (1997) 6.6.1-6.6.21
Copyright © 1997 by John Wiley & Sons, Inc.

BOSTON / CENTRAL UNIVERSITY LIBRARY

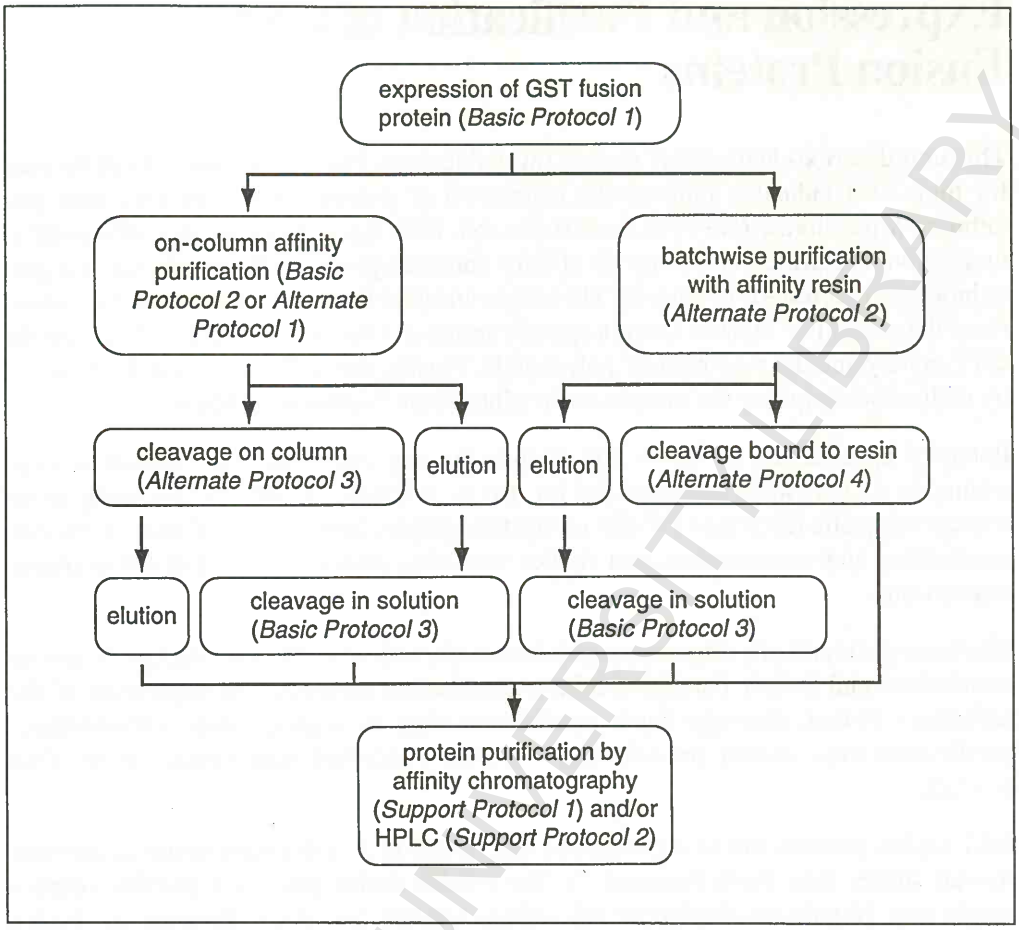

expression of GST fusion protein (*Basic Protocol 1*)
on-column affinity purification (*Basic Protocol 2* or *Alternate Protocol 1*)
batchwise purification with affinity resin (*Alternate Protocol 2*)
cleavage on column (*Alternate Protocol 3*)
elution
elution
cleavage bound to resin (*Alternate Protocol 4*)
elution
cleavage in solution (*Basic Protocol 3*)
cleavage in solution (*Basic Protocol 3*)
protein purification by affinity chromatography (*Support Protocol 1*) and/or HPLC (*Support Protocol 2*)

Figure 6.6.1 Flow chart showing the relationships between the various protocols in this unit.

fusion proteins using baculovirus systems (Davies et al., 1993) and yeast (Mitchell et al., 1993) have also been reported.

One factor that influences which pGEX vector to choose is whether or not the GST moiety will ultimately be cleaved away from the protein of interest. The pGEX-2T and pGEX-4T series of vectors contain a protease cleavage site for thrombin, and the pGEX-3X and pGEX-5X series of vectors contain protease cleavage sites for factor Xa. A more recently developed expression vector is the pGEX-6P series, which contains a cleavage site for PreScission protease (Pharmacia Biotech). PreScission protease has the advantage that it is effective at low temperatures (5°C). It is also a GST fusion protein, a feature that facilitates removal of the protease from the target protein after cleavage. Fusion proteins with a thrombin recognition site have the advantage that relatively small amounts of thrombin and short digestion times at 37°C will often cleave the fusion protein with high efficiency. Thrombin digestions are often the most cost effective on a per milligram of cleaved target polypeptide basis. Factor Xa is more expensive and typically requires use of much higher enzyme-to-substrate ratios for efficient cleavage. Solutions of factor Xa also have a more limited shelf life since freezing and thawing inactivates this enzyme.

For preparation of cDNA inserts encoding the desired polypeptide, see Ausubel et al. (1994) or Sambrook et al. (1989). Briefly, a set of oligonucleotides is designed for polymerase chain reaction (PCR) amplification of the region of interest of a pertinent cDNA. These oligonucleotides should also contain appropriate restriction sites adjacent to the desired coding region that are compatible with a restriction site in the cloning site of the selected pGEX vector (see Figs. 6.6.2 and 6.6.3). PCR amplification is performed,

ECU IAS/CERTAR UNIVERSITY LIBRARY

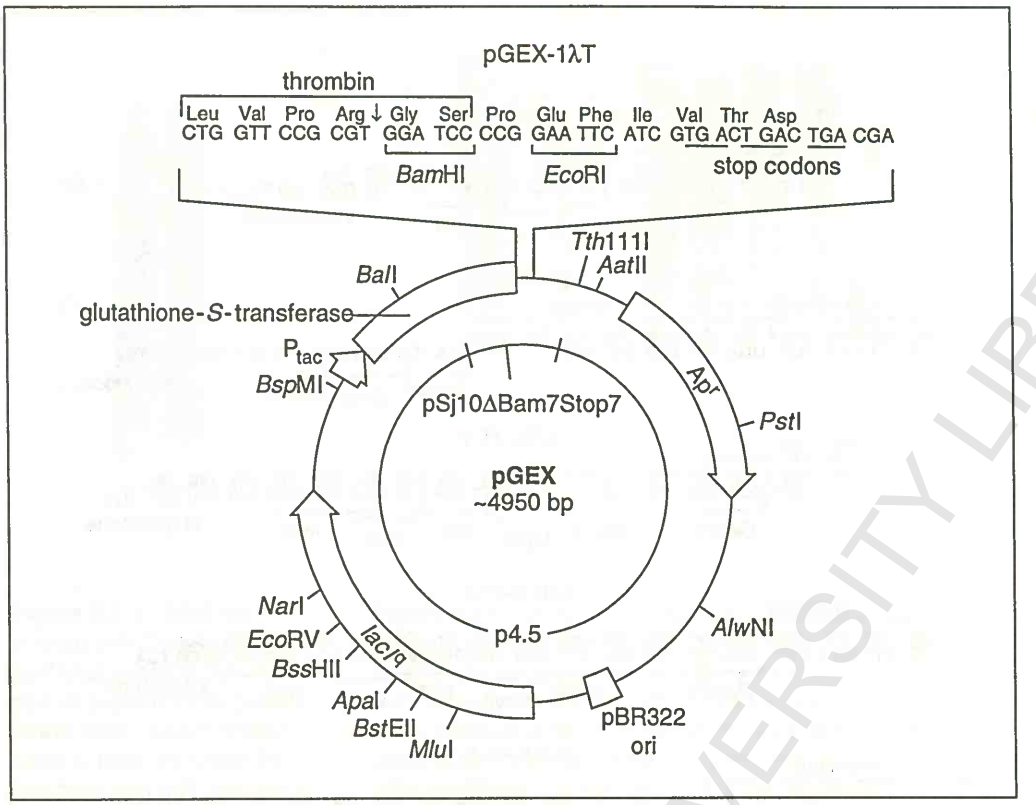


Figure 6.6.2 pGEX vectors are plasmid expression vectors that express a cloned gene as a fusion protein with glutathione-*S*-transferase (GST). The *lac* repressor gene binds to the *lac* promoter (p_{tac}) and represses expression of the GST fusion protein until induction with isopropyl-1-thio-β-D-galactopyranoside (IPTG). The polypeptide of interest can be inserted immediately after the GST gene using the polylinker site shown in brackets (pGEX-1λT, shown here, is the most common; see Fig. 6.6.3 for other PGEX polylinkers). Protease cleavage sites (brackets above the polylinker sequences) are located between the GST carrier protein and the protein of interest so that the GST moiety can be removed. Restriction endonuclease sites are indicated below the sequence of the polylinker and on the plasmid. An important consideration in selecting a vector and appropriate cloning sites is to minimize the number of extraneous residues introduced into the N-terminal of the target polypeptide. Vector map courtesy of Pharmacia Biotech.

followed by digestion of the PCR product and the pGEX vector with the appropriate restriction enzymes. The PCR product is then ligated into the pGEX vector and transfected into a suitable *E. coli* host. Several transformants should be grown in minicultures and induced with isopropyl-1-thio-β-D-galactopyranoside (IPTG) to check for expression of the desired fusion protein. Fusion protein expression can be monitored by SDS-PAGE (*UNIT 10.1*) or by Western blot (*UNIT 10.10*) detection of the GST fusion protein using an antibody specific for either the target protein or the GST moiety. Once successful expression is achieved, the integrity of the DNA should be verified by sequencing to ensure that no errors were introduced during PCR.

Before conducting a large-scale purification, it is worthwhile to perform a small pilot purification (~10-fold less than protocol descriptions) to determine optimal conditions. The purification can then easily be scaled up. All stages of purification should be monitored using SDS-PAGE (*UNIT 10.1*). In most cases, GST fusion protein expression is very high and a major band at the expected molecular weight (the GST moiety contributes 26 kDa to the molecular weight) is obvious when uninduced and induced cells are compared on SDS gels (Fig. 6.6.4). This band can then be monitored at each step of the purification. However, as noted above, if the level of protein expression obtained is low or band identification is ambiguous, the fusion protein can be monitored by Western blot

POSTMASTER THE UNIVERSITY LIBRARY

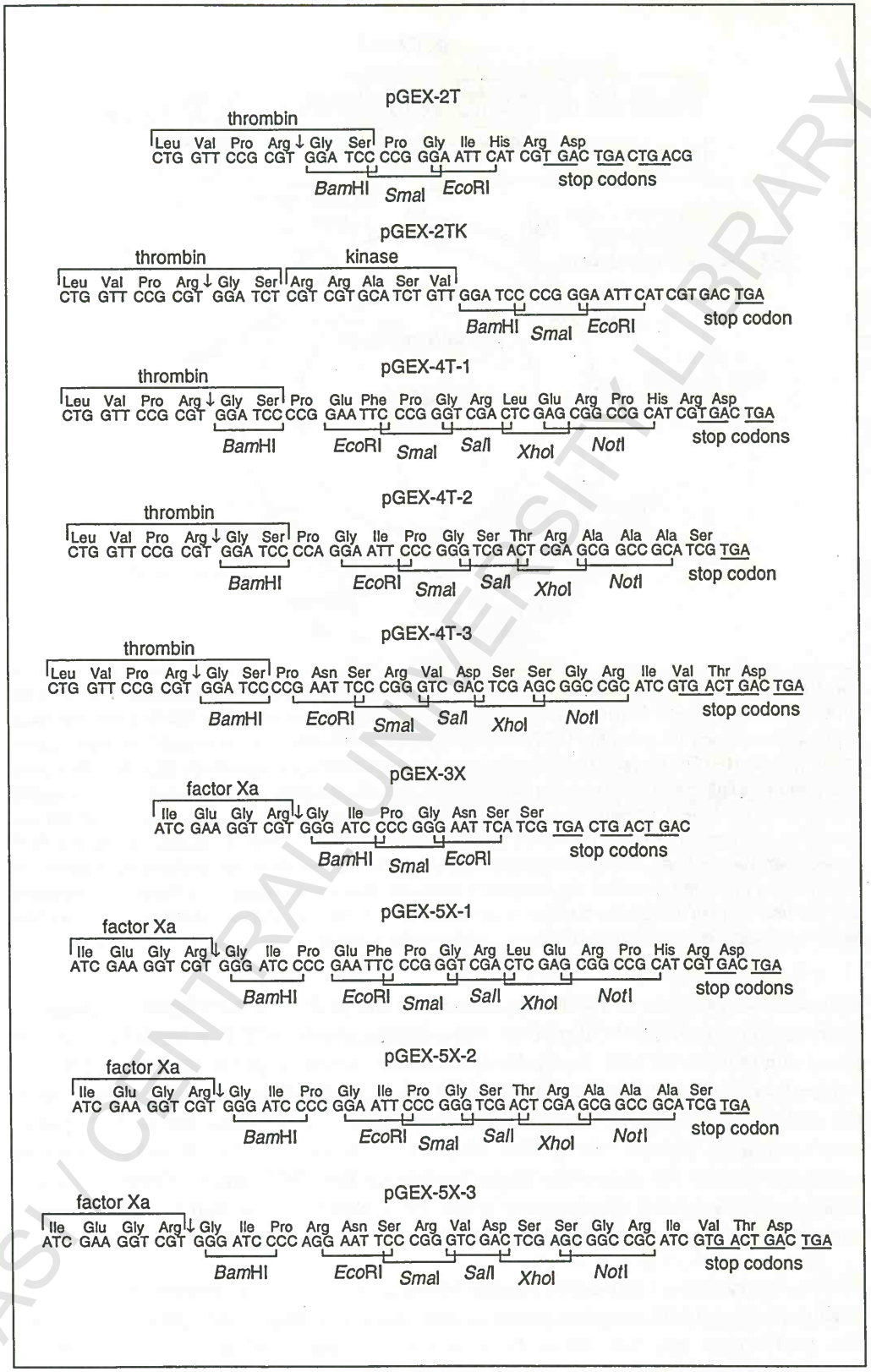

pGEX-2T
thrombin
Leu Val Pro Arg ↓ Gly Ser Pro Gly Ile His Arg Asp
CTG GTT CCG CGT GGA TCC CCG GGA ATT CAT CGT GAC TGA CTG ACG
BamHI SmaI EcoRI
stop codons
pGEX-2TK
thrombin
kinase
Leu Val Pro Arg ↓ Gly Ser Arg Arg Ala Ser Val
CTG GTT CCG CGT GGA TCT CGT CGT GCA TCT GTT GGA TCC CCG GGA ATT CAT CGT GAC TGA
BamHI SmaI EcoRI
stop codon
pGEX-4T-1
thrombin
Leu Val Pro Arg ↓ Gly Ser Pro Glu Phe Pro Gly Arg Leu Glu Arg Pro His Arg Asp
CTG GTT CCG CGT GGA TCC CCG GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT CGT GAC TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codons
pGEX-4T-2
thrombin
Leu Val Pro Arg ↓ Gly Ser Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
CTG GTT CCG CGT GGA TCC CCA GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codon
pGEX-4T-3
thrombin
Leu Val Pro Arg ↓ Gly Ser Pro Asn Ser Arg Val Asp Ser Ser Gly Arg Ile Val Thr Asp
CTG GTT CCG CGT GGA TCC CCG AAT TCC CGG GTC GAC TCG AGC GGC CGC ATC GTG ACT GAC TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codons
pGEX-3X
factor Xa
Ile Glu Gly Arg ↓ Gly Ile Pro Gly Asn Ser Ser
ATC GAA GGT CGT GGG ATC CCC GGG AAT TCA TCG TGA CTG ACT GAC
BamHI SmaI EcoRI
stop codons
pGEX-5X-1
factor Xa
Ile Glu Gly Arg ↓ Gly Ile Pro Glu Phe Pro Gly Arg Leu Glu Arg Pro His Arg Asp
ATC GAA GGT CGT GGG ATC CCC GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT CGT GAC TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codons
pGEX-5X-2
factor Xa
Ile Glu Gly Arg ↓ Gly Ile Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
ATC GAA GGT CGT GGG ATC CCC GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codon
pGEX-5X-3
factor Xa
Ile Glu Gly Arg ↓ Gly Ile Pro Arg Asn Ser Arg Val Asp Ser Ser Gly Arg Ile Val Thr Asp
ATC GAA GGT CGT GGG ATC CCC AGG AAT TCC CGG GTC GAC TCG AGC GGC CGC ATC GTG ACT GAC TGA
BamHI EcoRI SmaI SalI XhoI NotI
stop codons

Figure 6.6.3 Variations on the polylinker site shown in Figure 6.6.2 (courtesy of Pharmacia Biotech).

ECU IAS / CENTRAL UNIVERSITY LIBRARY

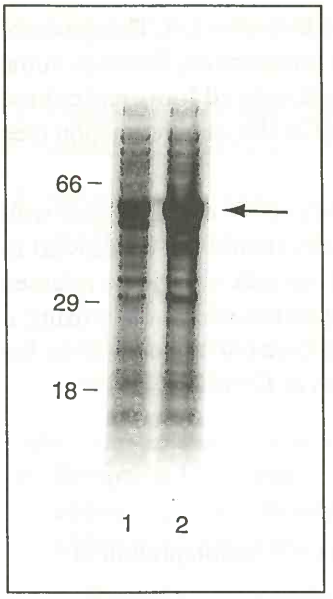


Figure 6.6.4 SDS gel stained with Coomassie brilliant blue showing the expression of a glutathione-*S*-transferase (GST) fusion protein. Lane 1, total cell lysate before induction with isopropyl-1-thio-β-D-galactopyranoside (IPTG). Lane 2, total cell lysate after 3-hr induction with IPTG. The position of the fusion protein is indicated by the arrow.

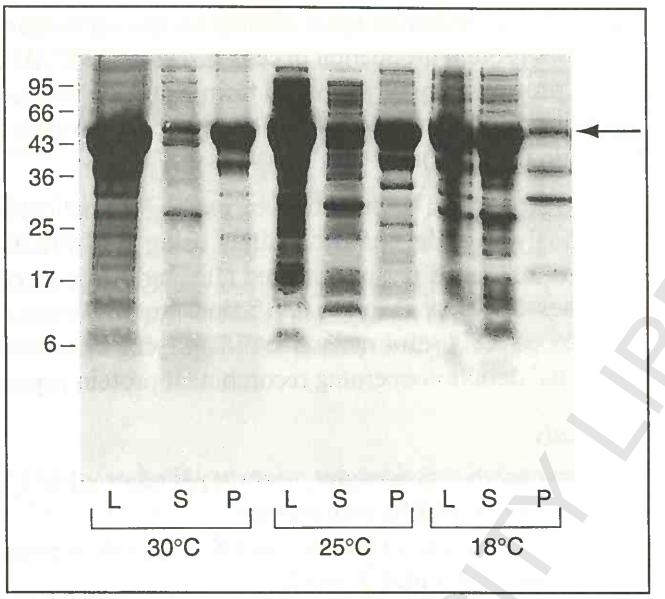


Figure 6.6.5 SDS gel stained with Coomassie brilliant blue showing the effects of different growth temperatures on solubility of an expressed glutathione-*S*-transferase (GST) fusion protein. *E. coli* transfected with a recombinant pGEX-2T vector were grown at different temperatures as indicated. Proportional aliquots of whole-cell lysate after sonication (L), the supernatant after centrifuging the lysate (S), and the remaining pellet (P) are shown. This recombinant was primarily in inclusion bodies (P) at 37°C (not shown), 30°C, and 25°C. At 18°C, however, nearly all of the expressed protein was in a soluble native form.

analysis (*UNIT 10.10*) using a GST-specific antibody (Pharmacia Biotech). It is recommended that the lysed cell extract, extracted pellet, and all other collected fractions from the purification be saved on ice until after careful analysis of the entire purification by SDS-PAGE and/or immunoblotting to ensure that fractions containing the fusion protein are not mistakenly discarded.

Basic Protocol 1 describes protein production in cells grown at 37°C; however, at this temperature some fusion proteins may be found in inclusion bodies in a denatured form. As an alternative to attempting to renature the protein after extraction from inclusion bodies (*UNIT 6.1*), expression at lower temperatures, such as 30°, 25°, 20°, or 15°C, can be evaluated to determine whether the protein can be obtained in good yield in the soluble fraction (Fig. 6.6.5). When expressing fusion proteins at lower temperatures, the initial overnight culture can be grown at 37°C followed by growth at a lower temperature prior to induction.

EXPRESSION OF GLUTATHIONE-*S*-TRANSFERASE FUSION PROTEIN

BASIC PROTOCOL 1

Transformed *E. coli* cells expressing the glutathione-*S*-tranferase (GST) fusion protein of interest are grown in culture in the presence of isopropyl-1-thio-β-D-galactopyranoside (IPTG) at the desired preparative scale. Since the expression level of GST fusion proteins is usually very high, adequate amounts of protein can usually be conveniently obtained by preparing a few liters of cells grown in shaker cultures. This protocol describes the preparation of 1.8 liters of transfected *E. coli* in three 600-ml units using 2-liter flasks in a shaker incubator. Moderate further scale-up is feasible by using more or larger flasks.

BOSTON UNIVERSITY LIBRARY

Further scale-up can be accomplished using a fermentor (see *UNITS 5.3 & 5.4*). This protocol describes protein production in cells grown at 37°C. At this temperature, however, some fusion proteins may be recovered from inclusion bodies in a denatured form, and culture conditions may need to be modified to improve protein yield in the soluble fraction (see Strategic Planning).

Culture growth can be monitored by reading the optical density at 550 nm (OD_{550}) as well as by analysis of the bacterial culture using SDS-PAGE. Cells should not be allowed to grow for extended periods of time after induction as cell lysis can occur; this releases proteases that may degrade the fusion protein. Visual inspection of the cells using a microscope is a useful method for identifying cell breakage (see *UNITS 5.1-5.3 & 6.1-6.5* for additional details concerning recombinant protein expression in *E. coli*).

Materials

Luria broth (LB medium; *UNIT 5.2*; pH adjusted to 7.2)
5 mg/ml ampicillin (see recipe)
Glycerol culture of transformed *E. coli* cells expressing GST fusion protein of interest in a pGEX vector
100 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG; see recipe)

2-liter culture flasks
500-ml culture flasks
Large centrifuge bottles (e.g., 1-liter capacity)
Low-speed refrigerated centrifuge (e.g., Beckman J6-B and JS-4.2 rotor or equivalent), 4°C

Additional reagents and equipment for SDS-polyacrylamide gel electrophoresis (SDS-PAGE; *UNIT 10.1*)

Grow bacterial cells

1. Prepare LB medium and add 600 ml to each of three 2-liter flasks and 100 ml to each of two 500-ml flasks. Autoclave 20 to 30 min at a slow exhaust (liquid) setting.

   Flasks should be filled to only 20% to 30% of their capacity to ensure adequate aeration of the medium during cell growth. Autoclave LB medium immediately after preparing it to prevent any incidental bacterial growth from occurring. It can then be stored up to 1 month at room temperature under sterile conditions.

2. Allow medium to cool to room temperature. Add 1 ml of 5 mg/ml ampicillin to one flask containing 100 ml LB medium.

   It is important to allow medium to cool before proceeding, because ampicillin is inactivated at temperatures >50°C. Flame the opening of all bottles and flasks to reduce the risk of contamination.

3. Using an inoculating loop, transfer some of the glycerol culture containing the transfected *E. coli* expressing the GST fusion protein of interest to the flask containing 100 ml LB medium with ampicillin.

   Sterile flame the inoculating loop as well as the opening of all bottles. Allow the loop to cool before transferring the inoculating culture. If the loop temperature is too high, all the cells could be killed during the transfer.

4. Incubate the inoculated culture on an environmental shaker set at 250 to 300 rpm overnight at 37°C.

5. The next morning, remove the culture from the shaker and read the optical density at 550 nm (OD_{550}) using a UV/visible light spectrophotometer.

   The OD_{550} of the overnight culture should be ~1.0. Use medium from the second 500-ml flask that did not receive ampicillin as a reference to zero the spectrophotometer.

ECU IASHC CENTRAL UNIVERSITY LIBRARY

6. Using sterile technique, add 6 ml of 5 mg/ml ampicillin to each 2-liter flask containing 600 ml LB medium (0.1 mM final concentration).

7. Dilute the overnight culture 1:20 by adding 30 ml to each of the three 2-liter flasks containing 600 ml LB medium.

8. Incubate the 600-ml cultures on a shaker at 37°C at 250 to 300 rpm until the OD_{550} is 0.5 to 0.7.

It should take ~2 hr for the culture to reach this early log stage of growth at 37°C. If cells are grown at lower temperatures to shift the expressed protein from inclusion bodies into the soluble fraction, this time must be increased, since the cells will grow more slowly at lower temperatures.

Induce expression of fusion protein

9. Remove a 1-ml sample from each flask and save for gel analysis. Induce cells by adding 6 ml of 100 mM IPTG per flask (1.0 mM final). Incubate at 37°C for an additional 2.5 to 3 hr.

Optimal growth conditions (OD_{550} at time of induction, growth temperature, and growth time after induction) for each recombinant protein should be empirically determined. The conditions given here will generally be suitable for growth at either 37° or 30°C. At lower temperatures (≤25°C), longer growth times will definitely be needed, and the best protein yields will usually be obtained if the OD_{550} prior to induction is 0.65 to 0.85.

10. Remove cultures from the shaker 2.5 to 3 hr after induction. Remove 1 ml from each culture and save for gel analysis. Check final OD_{550}.

Culture growth can be monitored at OD_{550}. When the cells reach saturation, they will stop dividing. A typical SDS-PAGE gel of an uninduced control culture and an induced culture after 3 hr growth is shown in Figure 6.6.4.

Recover bacterial cells

11. Pour each culture into a large centrifuge bottle and centrifuge 20 min at 4000 × *g*, 4°C.

Medium from the second 500-ml flask without ampicillin can be used to balance the centrifuge bottles.

12. Carefully decant the supernatant, leaving 15 to 50 ml in the centrifuge bottle. Resuspend the pelleted cells in the remaining supernatant. Transfer the cell suspension to a 50-ml centrifuge tube.

13. Centrifuge 20 min at 4000 × *g*, 4°C. Decant the supernatant.

14. Freeze cell pellet by placing in a –80°C freezer.

15. Analyze saved 1-ml samples from before and after induction using SDS-PAGE (*UNIT 10.1*) to check for protein expression.

As noted above, a 1-ml aliquot of culture is removed prior to induction and a second 1-ml sample is removed after the 3-hr induction. These aliquots can be centrifuged for 2 min and the supernatants carefully removed with a pipet. These samples can be stored at 0° to 4°C overnight prior to running a gel; for longer-term storage prior to gel analysis, store frozen (–20°C) to prevent proteolysis. The cell pellet can be directly resuspended in SDS sample buffer (200 µl) and heated 3 to 5 min at 90°C. Typical results are shown in Figure 6.6.4.

BCU LASCI CENTRAL UNIVERSITY LIBRARY

AFFINITY CHROMATOGRAPHY PURIFICATION OF A SOLUBLE GST FUSION PROTEIN

Soluble glutathione-*S*-transferase (GST) fusion proteins can easily be purified from cell lysate supernatants by affinity chromatography on glutathione–Sepharose 4B using either batch or column loading of the sample. Protease inhibitors should be added to the lysis buffer to minimize potential proteolysis. As an aid for cell lysis, 1% (v/v) Triton X-100 may be added to the lysis buffer. After cells are lysed, centrifugation is used to pellet any unlysed cells and inclusion bodies. Location of the fusion protein can usually easily be determined by analyzing a small aliquot of both the supernatant and the pellet by SDS-PAGE (*UNIT 10.1*). The glutathione–Sepharose 4B can be regenerated and reused for subsequent purifications. To avoid potential cross-contamination of different proteins or mutant forms of a single protein, however, it is recommended that a given column or batch of resin be reserved for use with a single protein. Preequilibration of the glutathione column (steps 1 to 4) can be performed either prior to or in parallel with cell lysis. All steps except SDS-PAGE (room temperature) should be performed in a cold room at 4°C, unless otherwise noted.

Materials

Glutathione–Sepharose 4B resin (Pharmacia Biotech)
PBS (*APPENDIX 2E*)
Glutathione buffer (see recipe)
PBS/EDTA/PMSF buffer (see recipe)
Pelleted *E. coli* culture expressing fusion protein (see Basic Protocol 1, step 14)
Lysis buffer (see recipe), ice cold
Wash buffer (see recipe), ice cold
PBS/EDTA (see recipe)

2.5 × 8–cm glutathione–Sepharose 4B column (e.g., Bio-Rad Econo)
Peristaltic pump
Sonicator equipped with microtip probe (e.g., Branson)
Dounce homogenizer
60-ml centrifuge bottles (capable of handling force of 48,000 × *g*)
High-speed refrigerated centrifuge (e.g., Beckman JZ-21M centrifuge and JA-18 rotor or equivalent), 4°C

Additional reagents and equipment for pouring chromatographic columns (*UNIT 8.3*) and for SDS-PAGE (*UNIT 10.1*)

Preequilibrate glutathione column

1. Pour 20 ml glutathione–Sepharose 4B resin into 2.5 × 8–cm column (glutathione column; see *UNIT 8.3* for details of pouring procedure).

   This amount of resin should be adequate for purification of fusion protein from three 600-ml cultures that contain ~20 to 40 mg fusion protein per culture. Although glutathione–Sepharose 4B has an advertised minimum binding capacity of 8 mg/ml resin, the actual capacity may be substantially different and should be empirically determined. The actual amount of resin used and column size can be increased or decreased depending on the amount of fusion protein to be purified.

2. Wash the glutathione column with 5 to 10 bed volumes PBS at a flow rate of 1.5 ml/min to remove the ethanol storage solution.

   A bed volume is one-half the amount of glutathione–Sepharose 4B that was added to the column as a 50% slurry.

DCU IASC CENTRAL UNIVERSITY LIBRARY

New columns should be prepared as described in steps 1 to 4. Previously used columns should be preequilibrated using steps 3 to 4 only, to ensure that the column is completely reduced and has maximal binding capacity.

Use of a peristaltic pump is recommended for convenient control of flow rates. Compression of the resin bed indicates that the column pressure is too high and that the flow rate should be lowered.

3. Wash the glutathione column with 3 to 5 bed volumes glutathione buffer at 1.5 ml/min.

Previously used columns may become partially oxidized on storage and should be preequilibrated (steps 3 to 4) within 24 hr before use.

4. Wash the glutathione column with 10 bed volumes PBS/EDTA/PMSF at 1.5 ml/min.

Lyse cells

5. Resuspend each pelleted 600-ml culture in 15 ml ice-cold lysis buffer.

Pellets should be resuspended in 25 to 50 µl buffer per milliliter of culture.

6. Sonicate the suspension using a probe-tip sonicator ten times for 10 sec each, with 1-min rests between sonications to lyse the cells. Save a sample (~100 µl) of the lysate for gel analysis and transfer remainder to a 60-ml centrifuge tube.

The cells are usually adequately lysed at the point when the suspension turns a slightly darker color and becomes clearer. To minimize proteolysis in the sample, it is essential to keep the cells on ice throughout the sonication procedure, and sonication should be performed in short bursts to minimize sample heating. Avoid excessive sonication, as this can lead to co-purification of E. coli host proteins along with the fusion protein of interest. Avoid frothing during sonication, which can denature the fusion protein.

7. Centrifuge the lysate 20 min at 48,000 × *g*, 4°C.

Unbroken cells, large cellular debris, and inclusion-body protein will be pelleted.

8. Decant the supernatant containing the soluble fusion protein into a clean 50-ml centrifuge tube.

9. Add a volume of ice-cold wash buffer equal to the volume of lysis buffer used in step 5 to the pellet. Use a dounce homogenizer to resuspend the pellet.

Pellets should be resuspended in 25 to 50 µl buffer per milliliter of culture.

10. Analyze the lysate, supernatant, and resuspended pellet using SDS-PAGE (*UNIT 10.1*) to verify that the fusion protein is in the supernatant.

If the fusion protein is in the supernatant, proceed to the next step. If the protein is in the pellet, it is necessary to purify the GST fusion proteins from inclusion bodies (see Alternate Protocol 2) or to start over, shifting the protein into the supernatant by growing the cultures at a lower temperature (see Troubleshooting and Fig. 6.6.5).

Load the column

11. Load the supernatant onto a preequilibrated glutathione column (see step 4). Collect fractions and analyze multiple fractions across the unbound peak by SDS-PAGE (*UNIT 10.1*) to verify that the fusion protein has bound and that column capacity was not exceeded.

For a 2.5-cm diameter column, a sample-loading flow rate of ≤0.1 ml/min is recommended to permit complete binding of the fusion protein to the resin. Faster flow rates may decrease the yield of bound fusion protein due to slow kinetics of association at 4°C, unless a large excess of resin is used. This step is most conveniently performed overnight; the flow rate

B. C. U. „M. Eminescu" IAŞI
CHIMIE

BCU IASI CENTRAL UNIVERSITY LIBRARY

can be adjusted so that most of the supernatant will be loaded by the next morning. Do not allow the column to run dry.

SDS-PAGE analysis of fractions collected during sample loading will reveal whether the fusion protein is bound to the column or is present in the unbound fractions. Absence of fusion protein in early unbound fractions combined with its appearance in late unbound fractions indicates that column capacity has been exceeded. If this condition is observed, reduce the protein load or increase the column size.

12. Wash the column with 5 to 10 bed volumes PBS/EDTA/PMSF at 1.5 ml/min.

13. Wash the column with 10 bed volumes PBS/EDTA at 1.5 ml/min to remove the PMSF.

If samples are to be cleaved with thrombin or factor Xa, any serine protease inhibitor (e.g., PMSF) must be removed from the sample before cleavage.

Elute the fusion protein

14. Elute the fusion protein from the column by washing the column with 5 bed volumes glutathione buffer.

A flow rate of 0.3 ml/min for a 2.5-cm column is recommended to elute the fusion protein in a minimal volume. Elution of the fusion protein can be monitored at A_{280} either with an online UV monitor or by reading the absorption of individual fractions.

15. Analyze the fractions by SDS-PAGE (*UNIT 10.1*) and pool fractions containing the GST fusion protein. Store at 0° to 4°C.

Fusion protein should typically be >90% pure at this point.

ALTERNATE PROTOCOL 1

AFFINITY CHROMATOGRAPHY PURIFICATION OF GST FUSION PROTEIN FROM INCLUSION BODIES

In some cases, fusion proteins are entirely or primarily located in a denatured, aggregated form in inclusion bodies. Glutathione-*S*-transferase (GST) fusion proteins can often be purified from inclusion bodies after solubilization in urea or another denaturant followed by renaturation by dialysis (*UNIT 6.3*). Other methods of solubilization include addition of detergents such as Sarkosyl (*N*-laurylsarosine; Grieco et al., 1992; Frangioni, 1992). After denaturation and renaturation, it is important to ensure that the protein has regained its native conformation and function (*UNITS 6.4 & 6.5*). Although denatured GST will not bind to the glutathione column and hence will not be recovered by this method, the possibility that the GST moiety has refolded when the fusion partner has not folded properly should also be considered. As an alternative to purifying proteins from inclusion bodies, growing the cells at a lower temperature will often shift the fusion protein into the supernatant while still producing ≥10 mg of fusion protein per liter of culture when pGEX vectors are used (see Strategic Planning and see Troubleshooting). All steps should be performed in a cold room at 4°C unless otherwise noted.

Additional Materials (also see Basic Protocol 2)

U buffer (see recipe)
Triton X-100
PBS/glycerol buffer (see recipe)
Low-speed refrigerated centrifuge (e.g., Beckman J6-B and JS-4.2 rotor or equivalent), 4°C

Additional reagents and equipment for dialysis (*APPENDIX 3B*)

1. Preequilibrate the glutathione column, lyse the cells, and separate the lysate pellet, which includes the inclusion bodies, and supernatant (see Basic Protocol 2, steps 1 to 9).

DCU IASC CENTRAL UNIVERSITY LIBRARY

2. Centrifuge washed pellet (see Basic Protocol 2, step 9) 20 min at 48,000 × *g*, 4°C. Decant the supernatant and resuspend the pellet in 12 ml freshly prepared U buffer per 600 ml original culture. Incubate 2 hr on ice.

   Pellets should be resuspended in 20 µl U buffer per ml of culture.

3. Centrifuge 20 min at 48,000 × *g*, 4°C. Carefully transfer the supernatant to a clean 50-ml centrifuge tube.

   The extracted fusion protein should now be in the supernatant.

4. Add Triton X-100 to the supernatant to give a final concentration of 1% (v/v).

5. Dialyze the sample (*APPENDIX 3B*) 2 to 3 hr in PBS/glycerol buffer.

   Dialysis buffer volume should be 20 times the sample volume.

6. Dialyze sample overnight in PBS/EDTA/PMSF.

   Dialysis buffer volume should be >100 times the sample volume.

7. Remove the sample from the dialysis bag and centrifuge 20 min at 4000 × *g*, 4°C.

8. Column purify the fusion protein (see Basic Protocol 2, steps 11 to 15).

BATCH PURIFICATION OF GST FUSION PROTEIN

ALTERNATE PROTOCOL 2

Soluble glutathione-*S*-transferase (GST) fusion proteins in cell lysate supernatants or renatured proteins extracted from inclusion bodies can be batch purified on glutathione–Sepharose 4B as an alternative to column-based purification (see Basic Protocol 2 and see Alternate Protocol 1). Batch purification requires less equipment and is relatively quick and easy, but resulting protein yield and sample purity are lower than in a chromatographic separation. In addition, the room temperature incubations recommended by the resin manufacturer (Pharmacia Biotech), especially the batch incubation of *E. coli* lysate supernatant or inclusion body extract with glutathione-Sepharose, increase the risk of proteolytic degradation of the fusion protein.

Prepare fusion protein

1. Extract soluble fusion proteins (see Basic Protocol 2, steps 5 to 10).

   If the fusion protein is in inclusion bodies, lyse the cells and collect the pellet (Basic Protocol 2, steps 5 to 9, followed by Alternate Protocol 1, steps 2 to 7).

Prepare resin slurry

2. Prepare a 50% slurry of glutathione–Sepharose 4B in PBS: For each milliliter of bed volume, centrifuge 1.33 ml of a 75% slurry of glutathione–Sepharose 4B for 5 min at 500 × *g*, room temperature. Discard the supernatant. Wash with 10 bed volumes PBS. Invert tube containing resin several times to mix, then centrifuge 5 min at 500 × *g*, room temperature, and remove supernatant. Add 1 ml PBS for each 1.33 ml of original slurry. Mix well before using.

Bind fusion protein to resin

3. Add 2 ml of 50% slurry of equilibrated glutathione-Sepharose to each 100 ml bacterial lysate supernatant. Incubate 30 min at room temperature with gentle agitation on a platform or orbital shaker.

4. Centrifuge the suspension 5 min at 500 × *g*, room temperature. Remove supernatant and save at 0° to 4°C for later analysis by SDS-PAGE (*UNIT 10.1*) to determine the efficiency of binding of the fusion protein to the resin.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

5. Wash the pellet with 10 bed vol PBS.

   A bed volume is one-half the volume of the 50% slurry that was added to the column.

6. Centrifuge the suspension 5 min at 500 × *g*, room temperature. Discard the supernatant.

7. Repeat the wash and centrifugation steps for a total of three washes with 10 bed volumes each of PBS.

Elute fusion protein

8. Elute the bound fusion protein by gently resuspending the sedimented resin in 1.0 ml glutathione buffer per milliliter resin bed volume. Incubate 10 min at room temperature with gentle agitation.

9. Centrifuge the suspension 5 min at 500 × *g*, room temperature. Transfer supernatant to a separate tube.

10. Repeat the elution and centrifugation (steps 8 to 9) a total of three times. Store at 0° and 4°C.

    The supernatants may be pooled into one tube or analyzed separately by SDS-PAGE (UNIT 10.1) to monitor for fusion protein content.

    The yield of fusion protein can be monitored by measuring the absorbance at 280 nm (A_{280}). The extinction coefficient will partially depend on the absorbance characteristics of the experimental component of the fusion protein. For the GST moiety alone, the concentration can be estimated using 1 A_{280} = ~0.6 mg/ml protein.

    As noted above, batch purification at room temperature increases the risk of proteolytic digestion of the target protein. To minimize such degradation, this procedure can alternatively be performed in a cold room at 4°C with incubation times in steps 3 and 8 increased 2- to 4-fold.

BASIC PROTOCOL 3

PROTEASE CLEAVAGE OF FUSION PROTEIN IN SOLUTION TO REMOVE GST AFFINITY TAG

The glutathione-*S*-transferase (GST) affinity tag is removed by cleaving with thrombin (pGEX-T vectors) or factor Xa (pGEX-X vectors). Conditions for optimal cleavage of each recombinant fusion protein must be empirically determined. Some of the parameters that can be varied include temperature, enzyme-to-substrate ratio, length of incubation, and buffer conditions. Proteolysis can usually be performed in the glutathione buffer used to elute the fusion protein from the column, either with or without addition of NaCl or Ca^{2+}. After proteolysis, the glutathione must be removed by dialysis prior to rechromatography on the glutathione column to separate the cleaved target protein from the GST moiety and uncleaved fusion protein.

Materials

Solution of affinity-purified fusion protein (see Basic Protocol 2, step 15, or Alternate Protocol 2, step 10)
Thrombin (Sigma) reconstituted in water at 0.5 U/µl or 1 µg/µl factor Xa (Boehringer Mannheim)
0.15 M PMSF in isopropanol
PBS/EDTA/PMSF buffer (see recipe)
Beckman J6-B centrifuge and JS-4.2 rotor (or equivalent), 4°C
Additional reagents and equipment for dialysis (*APPENDIX 3B*) and SDS-PAGE (*UNIT 10.1*)

ECU IAS / CENTRAL UNIVERSITY LIBRARY

1. To the solution of affinity-purified fusion protein, add an appropriate amount of thrombin (when using pGEX-T vectors) or factor Xa (when using pGEX-X vectors) per microgram of purified fusion protein and digest protein 2 to 8 hr in a shaking water bath at 37°C (thrombin) or 25°C (factor Xa).

   The appropriate amount of enzyme must first be empirically determined for each fusion protein by surveying a range of conditions in a pilot analytical proteolysis experiment. A convenient method is to digest 100 µg fusion protein per condition over a range of enzyme-to-substrate ratios and for differing incubation times. Typical digestion times range from 2 to 8 hr and typical enzyme-to-substrate-ratios are 1:100, 1:350, 1:1000, and 1:3000 (in units of enzyme per microgram fusion protein) at 37°C for bovine plasma thrombin, or 1:10, 1:25, 1:50, 1:100, 1:300 (microgram enzyme per microgram fusion protein) at 25°C for bovine plasma factor Xa. At the desired times, proteolysis is stopped by adding an aliquot of the sample to boiling SDS sample buffer, and samples are analyzed on an SDS minigel (UNIT 10.1; 2 µg/lane) to determine the best digestion conditions. If digestion is incomplete at the highest enzyme-to-substrate ratio tested with digestion times of 6 to 8 hr, consider reengineering the protease cleavage site by introducing a linker between the GST moiety and target polypeptide to decrease steric hindrance. In some cases, >10-fold improvement in cleavage efficiency can be achieved by adding as few as two glycines next to the thrombin site. The SDS gel in Figure 6.6.6 shows a typical thrombin digestion optimization experiment where the amount of enzyme has been varied.

2. Stop preparative digestion by adding a 1:500 dilution of 0.15 M PMSF stock solution. Incubate sample an additional 15 min at 37°C for thrombin or 30 min at 25°C for factor Xa to covalently inhibit the enzyme with the PMSF.

3. Dialyze the sample (APPENDIX 3B) twice versus 2 liters PBS/EDTA/PMSF for a minimum of 4 hr per buffer change at 4°C.

   Complete removal of glutathione is important if samples are to be rechromatographed on glutathione-Sepharose to remove the GST moiety and uncleaved fusion protein. Larger volumes of dialysate may be necessary depending on sample volume: e.g., if the sample volume is >40 ml, increase the dialysis buffer volume to 4 liters per change or use three changes of buffer. In addition, since glutathione equilibrates slowly during dialysis, when

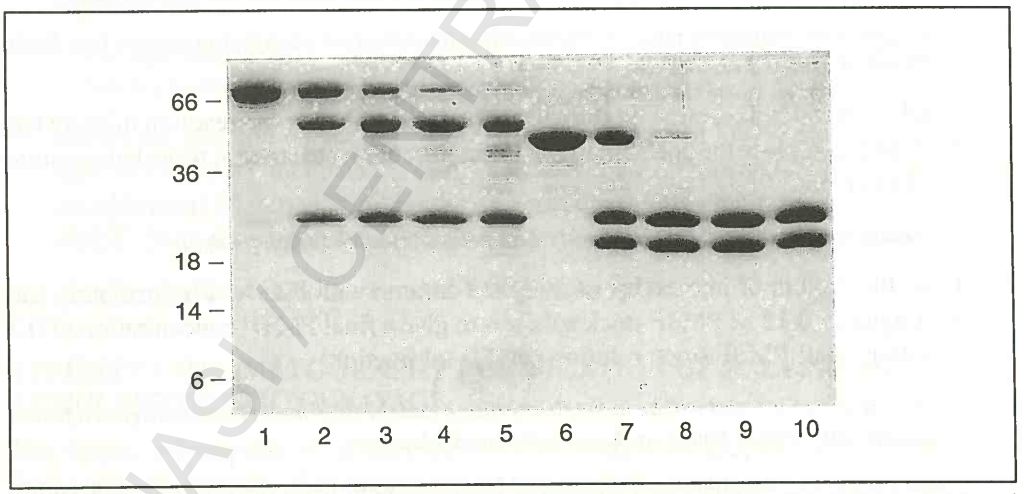


Figure 6.6.6 SDS gel stained with Coomassie brilliant blue showing pilot thrombin digestions of two recombinant glutathione-*S*-transferase (GST) fusion proteins. Samples were digested 3 hr at 37°C in buffer with varying enzyme-to-substrate ratios. Lane 1, a 78-kDa fusion protein. Lanes 2 to 5, thrombin digestion of the 78-kDa fusion protein using enzyme-to-substrate ratios of 1:3000, 1:1000, 1:350 and 1:100, respectively. Lane 6, a 45-kDa fusion protein. Lanes 7 to 10, thrombin digestion of the 45-kDa fusion protein using substrate ratios of 1:3000, 1:1000, 1:350, and 1:100, respectively. In each case, an enzyme-to-substrate ratio of 1:1000 was chosen as the optimal digestion condition

PGDIASI / CENTRAL UNIVERSITY LIBRARY

dialysis tubing with a MWCO <12,000 is used, the dialysis conditions should be increased to at least three changes with 2 liters dialysis buffer per change.

4. Centrifuge the dialyzed sample 20 min at 4000 × *g*, 4°C, to remove any precipitated material. Transfer the supernatant to a clean tube at 0° to 4°C and analyze by SDS-PAGE (*UNIT 10.1*).

The cleaved protein of interest can be separated from the GST moiety and uncleaved fusion protein by rechromatographing it on the glutathione column (see Support Protocol 1).

ALTERNATE PROTOCOL 3

PROTEASE CLEAVAGE OF FUSION PROTEIN BOUND TO A GLUTATHIONE-SEPHAROSE COLUMN

The target polypeptide can be cleaved from the glutathione-*S*-transferase (GST) moiety while the fusion protein is immobilized on the affinity matrix as described by Gearing et al. (1989). The fusion protein is loaded onto the glutathione column after cell lysis in the normal manner. Rather than elute the fusion protein with glutathione buffer, a solution of thrombin (or factor Xa) is loaded onto the column and allowed to incubate at room temperature for several hours. The cleaved protein is then eluted from the column by washing with PBS. This method is much faster and can yield higher purity products than batch protease digestions and in some cases, secondary cleavages can be avoided or minimized using the on-column method. Disadvantages of this method include much less effective protease cleavage, much lower final polypeptide yields, decreased control of digestion conditions, and inability to easily monitor the extent of cleavage using on column digestion.

Additional Materials (also see Basic Protocol 3)

PBS (*APPENDIX 2E*)
Glutathione–Sepharose 4B column containing bound fusion protein (see Basic Protocol 2, step 13)
Glutathione buffer (see recipe)

1. For each milliliter of glutathione–Sepharose 4B column containing bound fusion protein, dilute desired amount of reconstituted thrombin to 1.0 ml using PBS.

Enzyme concentrations must be empirically determined for each fusion protein (see Basic Protocol 3, step 1 annotation).

2. Load the thrombin reaction mixture onto the column. When the reaction mixture has been added, replace the bottom cap and allow the column to stand at room temperature for 2 to 16 hr.

Incubation times must be empirically determined for each fusion protein.

3. Elute the protein of interest by washing the column with PBS and immediately add an aliquot of 0.15 M PMSF stock solution to give a final PMSF concentration of 0.3 mM (e.g., 6 µl PMSF stock solution per 3.0-ml fraction).

The protein of interest will be in the flowthrough and the GST moiety and undigested fusion protein will remain bound to the glutathione-Sepharose.

4. Remove the GST moiety and uncleaved fusion protein from the column by washing with 5 bed volumes glutathione buffer at 0.3 ml/min.

5. Immediately after collecting fractions, store all fractions at 0° to 4°C.

6. Evaluate unbound and bound peaks using SDS-PAGE (*UNIT 10.1*).

A large proportion of the fusion protein is usually not cleaved by this method.

ECU IASI CENTRAL UNIVERSITY LIBRARY

BATCH PROTEASE CLEAVAGE AND SEPARATION OF FUSION PROTEIN ON GLUTATHIONE RESIN

ALTERNATE PROTOCOL 4

As a subsequent step to batch purification of soluble glutathione-*S*-transferase (GST) fusion proteins (see Alternate Protocol 2), the GST affinity tag can be cleaved from the protein of interest using thrombin (pGEX-2T) or factor Xa (pGEX-3X) before elution from the glutathione–Sepharose 4B resin in a manner analogous to the on-column cleavage described in Alternate Protocol 3. In this protocol, the fusion protein on the glutathione-Sepharose resin is incubated in the presence of protease, and the cleaved protein is released into the supernatant and collected by centrifugation.

Additional Materials (also see Basic Protocol 3)

Glutathione–Sepharose 4B resin with bound fusion protein (see Alternate Protocol 2, step 7)
PBS (*APPENDIX 2E*)

1. For each milliliter of glutathione–Sepharose 4B resin with bound fusion protein, dilute the desired amount of reconstituted thrombin to 1.0 ml with PBS.

   Enzyme concentrations must be empirically determined for each fusion protein (see Basic Protocol 3, step 1 annotation).

2. Add the reconstituted thrombin reaction mixture to the pelleted glutathione-Sepharose resin containing the bound fusion protein (from Alternate Protocol 2, step 7).

3. Gently resuspend the glutathione-Sepharose and agitate on an orbital shaker 2 to 16 hr at room temperature.

   Incubation times must be empirically determined for each fusion protein.

   Cleavage may be monitored by removing an aliquot of the slurry from the incubation mixture at different time points, centrifuging the slurry to separate the resin and supernatant, and analyzing the fractions by SDS-PAGE (UNIT 10.1).

4. Centrifuge the slurry 5 min at 500 × *g*, room temperature. Transfer supernatant to a separate tube.

   The supernatant will contain the protein of interest, while the GST moiety and uncleaved fusion protein will remain bound to the resin.

5. Inhibit the reaction by adding a 1:500 dilution of 0.15 M PMSF stock solution (0.3 mM PMSF final). Incubate the sample an additional 15 min at 37°C for thrombin or an additional 30 min at 25°C for factor Xa to covalently inhibit the enzyme with the PMSF. Store the samples at 0° to 4°C.

6. Analyze the final fractions by SDS-PAGE (*UNIT 10.1*).

AFFINITY CHROMATOGRAPHY PURIFICATION OF POLYPEPTIDES AFTER ENZYMATIC CLEAVAGE

SUPPORT PROTOCOL 1

The target polypeptide is affinity purified by rechromatography on a glutathione–Sepharose 4B column after enzymatic cleavage of the fusion protein and dialysis to remove glutathione present in the buffer from the initial column eluate. Uncleaved fusion protein and the glutathione-*S*-transferase (GST) moiety both bind to the column, and the protein of interest is in the column flowthrough. All steps should be performed in a cold room at 4°C unless otherwise indicated.

BOSTON / CENTRAL UNIVERSITY LIBRARY

Materials

Solution of fusion protein that has been cleaved with thrombin or factor Xa and dialyzed into PBS/EDTA/PMSF buffer (see Basic Protocol 3, step 4)
Glutathione buffer (see recipe)
PBS/EDTA/PMSF buffer (see recipe)
2.5 × 8–cm glutathione–Sepharose 4B column (e.g., Bio-Rad Econo)

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*)

1. Wash the glutathione–Sepharose 4B column with >3 bed volumes glutathione buffer at 1.5 ml/min.

   A bed volume is one-half the amount of glutathione-Sepharose that was added to the column.

   The same 2.5 × 8–cm column used for initial purification of a given fusion protein can be used for repurification of the same target polypeptide after protease cleavage. Glutathione-Sepharose must be fully reduced in order for the GST moiety to bind efficiently. If the column was washed with glutathione buffer <48 hr before, this step may be skipped.

2. Wash the glutathione column with 10 bed vol PBS/EDTA/PMSF at 1.5 ml/min.

3. Load the solution of fusion protein that has been cleaved onto the column using a flow rate of 0.1 ml/min. Collect fractions for later analysis by SDS-PAGE (*UNIT 10.1*) and store at 0° to 4° C.

   The protein of interest will be in the unbound or column flowthrough peak, and the GST moiety will bind to the glutathione-Sepharose. A low flow rate for sample application, such as 0.1 ml/min, should be used to ensure complete binding of GST to the column. Faster flow rates are likely to result in elution of excessive amounts of the GST moiety in the unbound fraction.

4. Wash the column with 2 to 3 bed volumes PBS/EDTA/PMSF at 1.5 ml/min.

5. Elute the bound GST moiety from the resin by washing with 5 bed volumes glutathione buffer at 0.3 ml/min.

6. Analyze bound and unbound peaks by SDS-PAGE (*UNIT 10.1*), pool the fractions containing the cleaved target polypeptide, and store at 0° to 4°C.

SUPPORT PROTOCOL 2

FURTHER PURIFICATION OF CLEAVED RECOMBINANT PROTEIN USING HPLC GEL FILTRATION

Cleaved proteins after rechromatography on glutathione columns (see Support Protocol 1) or eluted after on-column protease cleavage (see Alternate Protocol 3) contain the inactivated protease used to cleave the protein. These samples also typically contain small amounts of aggregates, a small amount of glutathione-*S*-transferase (GST) moiety that did not rebind to the column, and possibly other minor contaminants, including proteolytic cleavage products or host cell proteins. Although the protein is typically ~90% pure at this stage, a further purification step may be desirable for many applications. HPLC gel filtration is a fairly rapid high-resolution purification step that usually separates properly folded polypeptides from aggregates and incorrectly folded chains. Selection of appropriate matrices and column dimensions for particular applications is discussed in *UNIT 8.3*. Alternatively, ion-exchange chromatography (*UNIT 8.2*) can be used as a further purification strategy, either instead of or in addition to HPLC gel filtration. This protocol describes the HPLC gel filtration purification of cleaved polypeptides after rechromatography. All steps, including HPLC, should be performed at 4°C unless otherwise indicated.

ECU IAS / CENTRAL UNIVERSITY LIBRARY

Materials

Cleaved fusion protein (see Support Protocol 1, step 6, or Alternate Protocol 3, step 5)
PBS (*APPENDIX 2E*)

Centriprep concentrator with appropriate MWCO (Amicon)
Low-speed refrigerated centrifuge (e.g., Beckman J6-B and JS-4.2 rotor or equivalent, 4°C), *or* 0.22-µm low-protein-binding filter (Costar)
Gel-filtration columns (see *UNIT 8.3*)

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*)

1. Concentrate the unbound cleaved fusion protein using a Centriprep concentrator.

   Samples should be concentrated to a volume of no greater than 0.5% to 1% of the column volume to be used for the separation.

2. Centrifuge the concentrated sample 20 min at 4000 × *g*, 4°C. Remove the supernatant, being careful not to disturb any pellet that may have formed (precipitated material formed during concentration).

   Alternatively, a 0.22-µm low-protein-binding filter may be used to remove particulates. It is especially important to remove particulates to avoid clogging of the HPLC column end frits.

3. Inject the concentrated sample onto gel filtration column(s) equilibrated in PBS.
4. Monitor the absorbance at 280 nm (A_{280}) using an online HPLC detector and collect fractions.
5. Analyze the fractions by SDS-PAGE (*UNIT 10.1*) to determine fractions of interest and to evaluate the purity of the target polypeptide, which should be in the major peak.
6. Pool the desired fractions and store at 0° to 4°C.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or its equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Ampicillin, 5 mg/ml

500 mg ampicillin
100 ml H_2O
Filter sterilize using 0.22-µm filter
Store up to several months at 4°C

Do not autoclave the solution; ampicillin is inactivated above 50°C.

Glutathione buffer

50 mM Tris base
10 mM reduced glutathione (Sigma)
Adjust pH to 8.0 with 6 M HCl
Bring to final volume with cold H_2O
Prepare fresh daily and store at 4°C until needed

Isopropyl-1-thio-β-D-galactopyranoside (IPTG), 100 mM

1 g IPTG
42 ml sterile H_2O
Divide into 6-ml aliquots in sterile tubes
Store up to 1 year at –20°C

BCU IALH / CENTRAL UNIVERSITY LIBRARY

Lysis buffer

50 mM NaCl
50 mM Tris base
5 mM EDTA
1 µg/ml leupeptin
1 µg/ml pepstatin
0.15 mM phenylmethylsulfonyl fluoride (PMSF) in isopropanol
1 mM diisopropyl fluorophosphate (DFP)
Adjust pH to 8.0 with 6 M HCl
Bring to final volume with cold H_2O
Prepare fresh daily

CAUTION: *DFP is a dangerous neurotoxin. Handle the neat reagent with double gloves in a chemical fume hood only. Carefully follow all precautions supplied by the manufacturer for this chemical.*

PBS/EDTA

1× PBS (*APPENDIX 2E*)
5 mM EDTA
Adjust pH to 7.4 with 1 M NaOH
Bring to final volume with cold H_2O
Store up to 1 month at 4°C

PBS/EDTA/PMSF buffer

1× PBS (*APPENDIX 2E*)
5 mM EDTA
0.15 mM phenylmethylsulfonyl fluoride (PMSF) in isopropanol
Adjust pH to 7.4 with 1 M NaOH
Bring to final volume with cold H_2O
Store up to 1 week at 4°C

PBS/glycerol buffer

1× PBS (*APPENDIX 2E*)
20% (v/v) glycerol
1% (v/v) Triton X-100
5 mM 2-mercaptoethanol (2-ME)
5 mM EDTA
0.1 µg/ml leupeptin
0.1 µg/ml pepstatin
0.15 mM phenylmethylsulfonyl fluoride (PMSF) in isopropanol
0.1 mM diisopropyl fluorophosphate (DFP)
Adjust pH to 7.4 with 1 M NaOH
Bring to final volume with cold H_2O
Prepare fresh daily

CAUTION: *DFP is a dangerous neurotoxin. Handle the neat reagent with double gloves in a chemical fume hood only. Carefully follow all precautions supplied by the manufacturer for this chemical.*

U buffer

5 M urea
50 mM Tris base
5 mM EDTA
5 mM 2-mercaptoethanol (2-ME)
1 µg/ml leupeptin
1 µg/ml pepstatin

continued

DCU IASL/CENTRAL UNIVERSITY LIBRARY

0.15 mM phenylmethylsulfonyl fluoride (PMSF) in isopropanol
1 mM diisopropyl fluorophosphate (DFP)
Adjust pH to 8.0 with 6 M HCl
Bring to final volume with cold H_2O
Prepare fresh daily

CAUTION: *DFP is a dangerous neurotoxin. Handle the neat reagent with double gloves in a chemical fume hood only. Carefully follow all precautions supplied by the manufacturer for this chemical.*

Wash buffer

50 mM Tris base
5 mM EDTA
1 μg/ml leupeptin
1 μg/ml pepstatin
0.15 mM phenylmethylsulfonyl fluoride (PMSF) in isopropanol
Adjust pH to 8.0 with 6 M HCl
Bring to final volume with cold H_2O
Prepare fresh daily

COMMENTARY

Background Information

The pGEX vectors were designed for high-level, inducible, intracellular expression of glutathione-*S*-transferase (GST) fusion proteins produced in *Escherichia coli*. GST is a common 26-kDa protein of eukaryotes. The GST gene used in the development of the pGEX vectors was originally cloned from the parasitic helminth *Schistosoma japonicum* (Smith and Johnson, 1988).

Purification of fusion proteins from a whole-cell lysate is readily achieved through the strong affinity of the GST moiety for glutathione, which is immobilized on Sepharose beads. The fusion protein can be displaced under mild conditions from the glutathione-Sepharose beads using neutral-pH buffers containing free reduced glutathione.

The main advantages of this system are the very high level of fusion protein expression (often >10 to 50 mg/liter of *E. coli* culture grown on an environmental shaker) and the facile purification methods for both initial isolation and subsequent separation of cleaved polypeptide and GST moiety. Since nondenaturing purification conditions are employed, polypeptides that do not normally contain post-translational modifications usually retain their functional and antigenic properties. Other advantages of this system include availability of several alternative protease cleavage sites and the large number of bacterial hosts that can be used.

Purification of most soluble GST fusion proteins is straightforward, and success in the purification of insoluble products is determined largely by the ability to refold the fusion protein after extracting it from inclusion bodies. Since the GST moiety refolds fairly readily after urea solubilization, the presence of this moiety may facilitate refolding of the adjacent polypeptide. However, the high protein expression level of the vector, even at lower growth temperatures means that the preferred strategy for solubilizing proteins that are initially expressed in inclusion bodies is to attempt to grow the transfected cells at lower temperatures, where the protein often remains in the soluble fraction (see Fig. 6.6.5 and Troubleshooting). Difficulties with vector construction, protein expression, and protease cleavage may be greater when the molecular weight of the desired polypeptide is >100 kDa.

The GST moiety can usually be easily removed from the protein of interest by cleavage with thrombin or factor Xa. For some studies, such as some immunological or functional assays, removal of the GST moiety may not be required. An important consideration is that the GST moiety is a dimer; hence, steric hindrance effects and nonequivalence of functional sites on the associated fusion partner within this dimeric structure must be considered.

Troubleshooting

Contamination of the glutathione-*S*-transferase (GST) fusion protein after affinity purification with *E. coli* host cell proteins is usually an indication that sonication has been too severe. Other contaminants may represent de-

BOSTON UNIVERSITY LIBRARY

graded fragments of the fusion protein that might be eliminated by the addition of protease inhibitors and/or carefully keeping all samples, buffers, and sample tubes as cold as possible (i.e., keeping the materials at 0°C, on ice, often results in noticeably lower proteolysis than occurs where samples are kept at 4°C in a refrigerator). Degradation of fusion proteins during expression can sometimes be minimized by adding isopropyl-1-thio-β-D-galactopyranoside (IPTG) later in the bacterial culture growth with cells at higher OD at time of induction and by decreasing the duration of the induction period.

If the fusion protein does not completely bind to the glutathione-Sepharose resin, several options may be tried: increasing the quantity of glutathione-Sepharose, decreasing the protein loading flow rate, ensuring that the column is completely reduced by pretreating it with a freshly prepared glutathione buffer within 24 hr prior to using the column for purification, and ensuring that any denaturants or reducing reagents used to treat the sample have been completely removed through exhaustive dialysis.

Expression of GST fusion protein in inclusion bodies can be addressed in several ways. The preferred method is to lower the growth temperature from 37° to 30°, 25°, 20°, or 15°C. The optimal growth temperature is that temperature where >50% of the expressed protein is in the supernatant. This strategy is efficient even if the expressed fusion protein yield decreases somewhat at the lower growth temperatures. It is usually more time and cost effective to grow a larger quantity of cells than to attempt to renature fusion protein extracted from inclusion bodies since the latter method nearly doubles the time required to purify the fusion protein. In one study, when the authors' laboratory expressed eleven GST fusion proteins at 37°C, four proteins were primarily in inclusion bodies and the remaining seven proteins were partially in the supernatant, although a substantial amount of fusion protein was also in inclusion bodies. When these eleven fusion proteins were expressed at 30°C, ten were primarily (>80%) in the soluble fraction. The remaining recombinant protein was primarily in inclusion bodies at either 30° or 25°C, but could be switched to the soluble fraction and produced in high yield by growing the cells at 18°C, as shown in Figure 6.6.5. If lowering the growth temperature does not result in expression of a soluble protein, renaturing the protein after extracting inclusion bodies with a chaotropic reagent such as urea can be attempted (see Alternate Protocol 1). Finally, even when most of a protein is primarily in inclusion bodies, there may be enough protein in the soluble fraction that it can be purified in low milligram yields per liter of culture.

In some cases, using high enzyme-to-substrate ratios and prolonged incubation times is not sufficient to cleave the fusion protein in good yield. To minimize steric hindrance at the thrombin cleavage site, modified pGEX vectors that contain a glycine-rich, nine-amino-acid "kinker" region immediately before or after the thrombin cleavage site (Guan and Dixon, 1991; Hakes and Dixon, 1991) can be tried. Sometimes introduction of as few as one or two extra glycines immediately after to the thrombin site may substantially improve the rate of protease cleavage.

The probability that extraneous residues on the N-termini of cleaved polypeptides could significantly affect structural and functional properties of expressed polypeptides is very likely to increase as the length of the extraneous sequence increases. Hence, the number of extra residues remaining on the target polypeptide after protease cleavage should be minimized. Although thrombin cleavage is usually more cost effective than factor Xa cleavage, a minimum of two extra residues from the thrombin recognition site (GS) are left on the cleaved peptide. Although no residues from the factor Xa recognition site are left on the cleaved polypeptide, at least several extraneous amino acids are introduced onto the N-terminal of all pGEX-X vectors from the sequence remaining at the cloning site (see Figs. 6.6.2 and 6.6.3). Depending upon the pGEX vector and cloning site utilized, as few as two and as many as thirteen extra amino acids may remain on the cleaved N-terminal end of the target protein.

The yield of GST fusion proteins is typically quite high, i.e., 10 to 50 mg of fusion protein per liter of culture when cells are grown on an environmental shaker. When low protein expression does occur, it can often be improved by optimizing the bacterial culture growth conditions: for example, growing the cells to a somewhat higher OD_{550} at the time of induction, increasing the length of the induction period, using another strain of *E. coli* such as a protease-negative strain, or growing the cells at higher temperatures (although proteins are more likely to be in inclusion bodies at higher temperatures; see Background Information). Caution should be taken when using higher cell ODs and longer growth times after induction,

ECU IAS/CERIAS UNIVERSITY LIBRARY

since the cells may start to lyse under these conditions.

For additional information concerning GST fusion protein expression, Pharmacia Biotech publishes a reference list, bulletin number 18-1117-27, that classifies related publications into several useful categories to assist in troubleshooting a variety of parameters.

Anticipated Results

Yields of fusion protein can vary widely. Typical yields are 10 to 50 mg/liter, but can occasionally be much lower, especially if the fusion protein is toxic to the cells or is unstable. In some cases, >50 mg/liter can be obtained when expression conditions have been well optimized. A single-step affinity purification should yield fusion protein that is >90% pure in most cases. The relationship between yield of fusion protein and yield of cleaved, repurified recombinant target polypeptide is in part due to the mass of the target polypeptide. A good final yield of a cleaved protein after repurification on a glutathione column followed by HPLC gel filtration might be ~2 mg/liter for a 10-kDa protein and ~10 mg/liter for a 50-kDa protein.

Time Considerations

Protein expression takes 1.5 days of intermittent work requiring approximately 3 to 4 hr of operator time. Longer induction periods may be required at temperatures lower than 30°C, but total operator time remains the same. A small-scale batch purification can be completed in 1 day. A large-scale column purification, cleavage of fusion protein, and repurification of the cleaved peptide by affinity and gel filtration chromatography will take ~5 to 8 days of intermittent work requiring several hours of operator time per day. The purification should be completed in as short a time as practical to minimize proteolysis, aggregation, and precipitation of impure fractions.

Literature Cited

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.). 1994. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Davies, A.H., Jowett, J.B.M., Jones, I.M. 1993. Recombinant baculovirus vectors expressing glutathione-*S*-transferase fusion proteins. *Bio/Technology* 11:933-936.

Frangioni, J.V. 1992. Solubilization and purification of enzymatically active glutathione-*S*-transferase (pGEX) fusion proteins. *Anal. Biochem.* 210:179-187.

Gearing, D.P., Nicola, N.A., Metcalf, D., Foote, S., Willson, T.A., Gough, N.M., and Williams, R.L. 1989. Production of leukemia inhibitory factor in *Escherichia coli* by a novel procedure and its use in maintaining embryonic stem cells in culture. *Bio/Technology* 7:1157-1161.

Grieco, F., Hull, J., and Hull, R. 1992. An improved procedure for the purification of protein fused with glutathione-*S*-transferase. *Biotechniques* 13:856-857.

Guan, K.L. and Dixon, J.E. 1991. Eukaryotic proteins expressed in *Escherichia coli*: An improved thrombin cleavage and purification procedure of fusion proteins with glutathione-*S*-transferase. *Anal. Biochem.* 192:262-267.

Hakes, D.J. and Dixon, J.E. 1991. New vectors for high level expression of recombinant proteins in bacteria. *Anal. Biochem.* 202:293-298.

Mitchell, D.A., Marshall, T.K., and Deschenes, R.J. 1993. Vectors for the overexpression of glutathione-*S*-transferase fusion proteins in yeast. *Yeast* 9:715-722.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-*S*-transferase. *Gene* 67:31-40.

Key Reference

Smith and Johnson, 1988. See above.

Original description of the pGEX system.

Contributed by Sandra Harper and David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Expression and Purification of Thioredoxin Fusion Proteins

This unit describes a gene fusion expression system that uses thioredoxin, the product of the *Escherichia coli trxA* gene, as the fusion partner. The system is particularly useful for high-level production of soluble fusion proteins in the *E. coli* cytoplasm; in many cases heterologous proteins produced as thioredoxin fusion proteins are correctly folded and display full biological activity. Although the thioredoxin gene fusion system is routinely used for protein production, high-level production of peptides—i.e., for use as antigens—is also possible because the prominent thioredoxin active-site loop is a very permissive site for the introduction of short amino acid sequences (10 to 30 residues in length). The inherent thermal stability of thioredoxin and its susceptibility to quantitative release from the *E. coli* cytoplasm by osmotic shock can also be exploited as useful tools for thioredoxin fusion protein purification. In addition, a more generic method for purification of any soluble thioredoxin fusion employs a modified form of thioredoxin (called "His-patch Trx"), which has been designed to bind to metal chelate resins. Protein fusions to His-patch Trx can usually be purified in a single step from cell lysates (see Strategic Planning).

The first step is construction of a fusion of *trxA* to any desired gene and expression of the fusion protein in an appropriate host strain at 37°C (see Basic Protocol). Additional protocols describe *E. coli* cell lysis using a French pressure cell and fractionation (see Support Protocol 1), osmotic release of thioredoxin fusion proteins from the *E. coli* cytoplasm (see Support Protocol 2), and heat treatment to purify some thioredoxin fusion proteins (see Support Protocol 3).

STRATEGIC PLANNING

The thioredoxin gene fusion expression vectors pTRXFUS and hpTRXFUS, both of which carry the *E. coli trxA* gene (Fig. 6.7.1), are used for high-level production of C-terminal fusions to thioredoxin. The vector hpTRXFUS differs from pTRXFUS in that it contains a modified *E. coli trxA* gene which produces a mutant protein ("His-patch" thioredoxin) that can specifically bind to metal chelate matrices charged with nickel or cobalt, otherwise known as native metal-chelate affinity chromatography (MCAC; *UNIT 9.4*). The *trxA* translation-termination codon has been replaced in both vectors by DNA encoding a ten-residue peptide linker sequence that includes an enterokinase (enteropeptidase; LaVallie et al., 1993a) cleavage site. This highly specific site can be cleaved with enterokinase following purification of the fusion protein to release the protein of interest from its thioredoxin fusion partner. Immediately downstream of the DNA encoding the enterokinase site in pTRXFUS and hpTRXFUS lies a DNA polylinker sequence containing a number of unique restriction endonuclease sites that can be used for forming in-frame translational fusions of any desired gene to *trxA*. Downstream of the DNA polylinker lies the *E. coli aspA* transcription terminator. Replication of these vectors is controlled by a modified colE1 replication origin similar to that found in pUC vectors (Norrander et al., 1983). Plasmid selection and maintenance is ensured by the presence of the β-lactamase gene on the vector. The vector pALtrxA-781 (Fig. 6.7.1) is very similar to pTRXFUS. However in this plasmid the *trxA* gene is followed by a translation termination codon, and the sequences encoding the enterokinase-site peptide linker are absent. A unique *Rsr*II site, present in both pALtrxA-781 and pTRXFUS, allows for the easy insertion of short peptide-encoding DNA sequences into *trxA* within the region that encodes the active-site loop.

Contributed by John McCoy and Edward LaVallie

Current Protocols in Protein Science (1997) 6.7.1-6.7.14

Copyright © 1997 by John Wiley & Sons, Inc.

BRADFORD UNIVERSITY LIBRARY

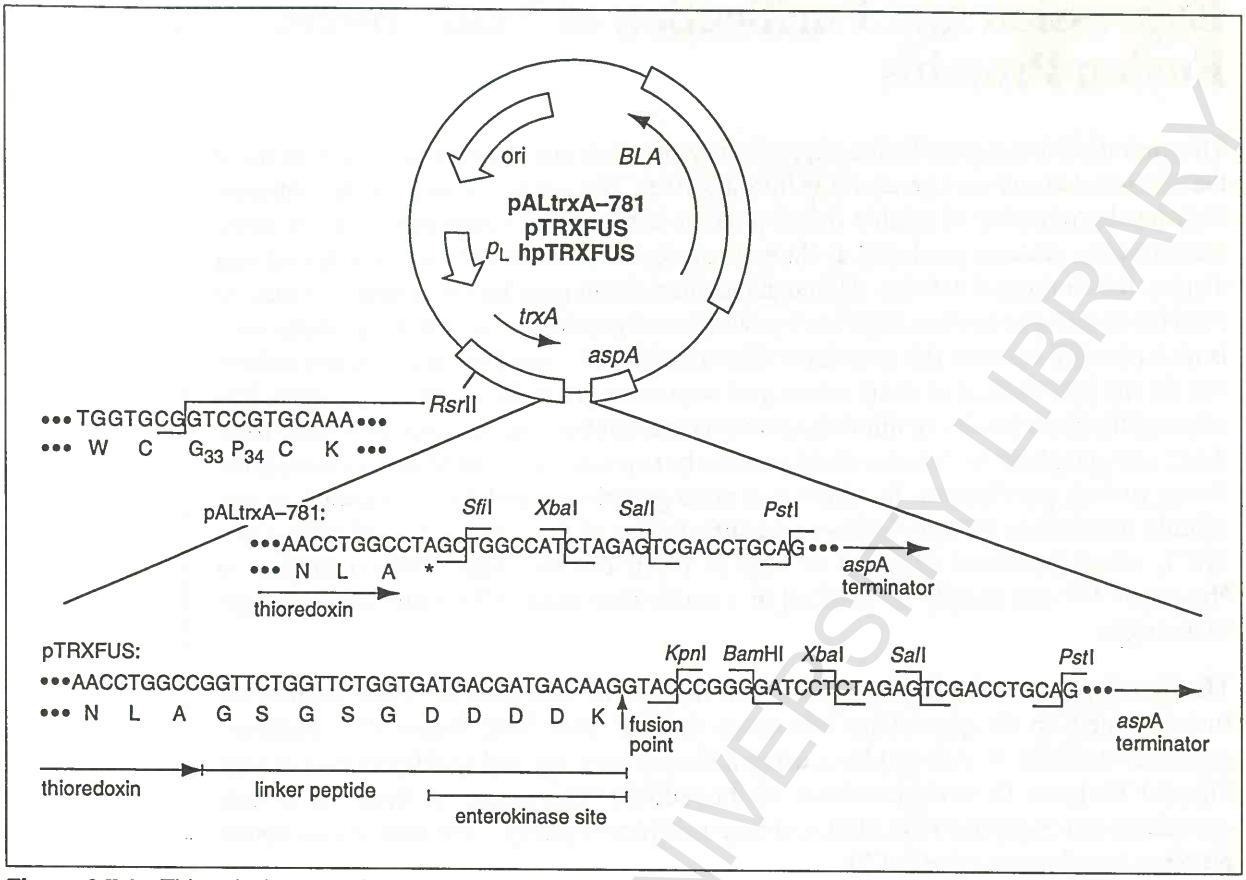


Figure 6.7.1 Thioredoxin gene fusion expression vectors pTRXFUS, hpTRXFUS, and pALtrxA-781. pALtrxA-781 contains a polylinker sequence at the 3′ end of the *trxA* gene. pTRXFUS and hpTRXFUS contain a linker region encoding a peptide that includes the enterokinase cleavage site between the *trxA* gene and the polylinker. The sequence surrounding the active site loop of thioredoxin has a single *Rsr*II site that can be used to insert peptide coding sequence. The asterisk indicates a translational stop codon. Abbreviations: *trxA*, *E. coli* thioredoxin gene; *BLA*, β-lactamase gene; ori, colE1 replication origin; pL, bacteriophage λ major leftward promoter; *aspA* terminator, *E. coli* aspartate amino-transferase transcription terminator.

pTRXFUS, hpTRXFUS, and pALtrxA-781 carry the strong bacteriophage λ promoter pL (Shimatake and Rosenberg, 1981) positioned upstream of the *trxA* gene. Transcription initiation at the pL promoter is controlled by the intracellular concentration of λ repressor protein (cI). cI^{857}-containing strains (Shatzman et al., 1990) can be used for heat inductions of pL at 42°C; alternatively, in the strains carrying the wild-type repressor, pL can be induced by a prior induction of the *E. coli* SOS stress response. However, it is often desirable to express heterologous genes in *E. coli* at temperatures considerably lower than 42°C, or under conditions where cells are not undergoing a physiological stress. Strains GI698, GI724 and GI723 were designed to allow the growth and induction of pL expression vectors, including pTRXFUS, hpTRXFUS, and pALtrxA-781, under mild conditions over a wide range of temperatures (see Table 6.7.1; Mieschendahl et al., 1986). Each of these strains carries a wild-type allele of cI stably integrated into the *E. coli* chromosome at the nonessential *ampC* locus. A synthetic *trp* promoter integrated into *ampC* upstream of the cI gene in each strain directs the synthesis of cI repressor only when intracellular tryptophan levels are low. When tryptophan levels are high, synthesis of cI is switched off; therefore, the presence of tryptophan in the growth medium of GI698, GI723, or GI724 will block expression of λ repressor and thus will turn on pL. Because the three strains carry ribosome-binding sequences of different strengths at the 5′-end of

PCU IASRI CENTRAL UNIVERSITY LIBRARY

Table 6.7.1 *E. coli* Strains for Production of Thioredoxin Fusion Proteins at Varying Temperatures

Strain	Desired production temperature (°C)	Pre-induction growth temperature (°C)	Induction period (hr)
GI698	15	25	20
GI698	20	25	18
GI698	25	25	10
GI724	30	30	6
GI724	37	30	4
GI723	37	37	5

their respective cI genes, they maintain intracellular concentrations of λ repressor that increase in the order GI698 < GI724 < GI723. The choice of which strain to use for a particular application is dependent on the desired culture conditions as described below.

Although some thioredoxin fusion proteins produced at 37°C are insoluble, expression at lower temperatures can often result in the fusion protein being produced in a soluble form. Each of the three pL host strains GI698, GI723, and GI724 is suitable for the production of thioredoxin fusion proteins over a particular temperature range. Table 6.7.1 indicates the correct strain for expression of thioredoxin fusion proteins at any temperature between 15°C and 37°C. The induction protocol at any of these temperatures is the same as that described for induction of GI724 at 37°C (see Basic Protocol), except the preinduction growth temperature and the length of the induction period vary according to the strain used and the temperature chosen. Cultures should be grown at the indicated preinduction growth temperature until they reach a density of 0.4 to 0.6 OD_{550}/ml. They should then be moved to the desired induction temperature and induced by the addition of 100 μg/ml tryptophan.

Low-temperature inductions are best performed in strain GI698. However, this strain makes only enough cI repressor protein to maintain the vectors in an uninduced state at temperatures below 25°C. GI698 should therefore never be grown above 25°C when it carries a pL plasmid. A nonrefrigerated water bath can be maintained below room temperature by placing it in a 4°C room and setting the thermostat to the desired temperature.

It is often a good idea to collect time points during the course of a long induction period and to fractionate cells from these time points (see Support Protocol 1, steps 9 to 13). Although a particular fusion protein may be soluble during the early part of an induction, during the later phases of induction it may become unstable or its concentration inside the cell may exceed a critical threshold above which it will precipitate and appear in the insoluble fraction.

BCU IAS CENTRAL UNIVERSITY LIBRARY

CONSTRUCTION AND EXPRESSION OF A THIOREDOXIN FUSION PROTEIN

This protocol describes construction and subsequent expression of a gene fusion between *trxA* (encoding thioredoxin) and a gene encoding a particular protein or peptide. After a clone carrying the correct fusion sequence is constructed, analyzed, and isolated, cultures are grown and expression is induced. The protocol is described in terms of the *E. coli* host strain GI724 with expression at 30°C; it may also be applied to strains GI698 and GI723 (also available from Genetics Institute) for expression at other temperatures by using the parameters specified in Table 6.7.1 (see Strategic Planning).

Materials

DNA fragment encoding desired sequence
Thioredoxin expression vectors (Fig. 6.7.1): pTRXFUS or pALtrxA-781 (Genetics Institute or Invitrogen) or hpTRXFUS (Genetics Institute)
E. coli strain GI724 (Genetics Institute or Invitrogen), grown in LB medium and made competent
LB medium (*UNIT 5.2*)
IMC plates (see recipe) containing 100 µg/ml ampicillin
CAA/glycerol/ampicillin 100 medium (see recipe)
IMC medium (see recipe) containing 100 µg/ml ampicillin
10 mg/ml tryptophan (see recipe)
SDS-PAGE sample buffer (see recipe)

30°C convection incubator
18 × 50–mm culture tubes
Roller drum (New Brunswick Scientific)
250-ml culture flask
70°C water bath
Microcentrifuge, 4°C

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*), and Coomassie brilliant blue staining (*UNIT 10.5*)

Construct the trxA gene fusion

1. Use DNA fragment encoding the desired sequence to construct either an in-frame fusion to the 3′-end of the *trxA* gene in pTRXFUS or hpTRXFUS, or a short peptide insertion into the unique *Rsr*II site of pALtrxA-781.

   A precise fusion of the desired gene to the enterokinase linker sequence in pTRXFUS or hpTRXFUS can be made by using the unique KpnI site trimmed to a blunt end with the Klenow fragment of E. coli DNA polymerase. The desired gene can usually be adapted to this blunt-end construct by using a synthetic oligonucleotide duplex ligated between it and any convenient downstream restriction site close to the 5′ end of the gene. When designing the fusion junction, note that enterokinase is able to cleave —DDDDK↓X—, where X is any amino acid residue except proline. Synthetic oligonucleotides encoding short peptides for insertion into the thioredoxin active-site loop at the RsrII site will insert only in the desired orientation, because the RsrII sticky end consists of three bases.

2. Transform the ligation mixture containing the new thioredoxin fusion plasmid into competent GI724 cells. Plate transformed cells onto IMC plates containing 100 µg/ml ampicillin to select transformants. Incubate plates in a 30°C convection incubator until colonies appear.

   Strains GI698, GI723, and GI724 are all healthy prototrophs that can grow under a wide variety of growth conditions, including rich and minimal media and a broad range of growth temperatures (see Table 6.7.1). These strains can be prepared for transformation with pL-containing vectors by growing them in LB medium at 37°C. LB medium may also

PCU IASI / CENTRAL UNIVERSITY LIBRARY

be used for these strains during the short period of outgrowth immediately following transformation. This growth period of 30 min to 1 hr is often used to express drug-resistance phenotypes before plating out plasmid transformations onto solid medium. Subsequently, however, these strains should be grown only on minimal or tryptophan-free rich media, such as IMC medium containing 100 µg/ml ampicillin (for expression of the fusion protein) or CAA/glycerol/ampicillin 100 medium (for plasmid DNA preparations). Except during transformation, LB medium should never be used with these three strains when they carry pL plasmids because LB contains tryptophan. The pL promoter is extremely strong and should be maintained in an uninduced state until needed so that expression of the protein will not lead to selection of mutant or variant cells with lower expression due to undesirable genetic selections or rearrangements in the expression strain.

3. Grow candidate colonies in 5 ml CAA/glycerol/ampicillin 100 medium overnight at 30°C. Prepare minipreps of plasmid DNA and check for correct gene insertion into pTRXFUS by restriction mapping.

4. Sequence plasmid DNA of candidate clones to verify the junction region between thioredoxin and the gene or sequence of interest.

Induce expression

5. Streak out frozen stock culture of GI724 containing thioredoxin expression plasmid to single colonies on IMC plates containing 100 µg/ml ampicillin. Grow 20 hr at 30°C.

   Occasionally there is induction of pL plasmids grown in GI698 and GI724 at 37°C, even in medium containing no tryptophan. Such induction appears to be a temperature-dependent phenomenon. If growth at 37°C prior to pL induction is essential, then GI723 should be used as the host strain because GI723 produces higher levels of cI repressor than both GI698 and to GI724. Otherwise, plasmid-containing GI698 should be grown at 25°C and plasmid-containing GI724 should be grown at 30°C prior to induction (see Table 6.7.1).

6. Pick a single fresh, well-isolated, colony from the plate and use it to inoculate 5 ml IMC medium containing 100 mg/ml ampicillin in an 18 × 150–mm culture tube. Incubate overnight at 30°C on a roller drum.

7. Add 0.5 ml overnight culture to 50 ml fresh IMC medium containing 100 µg/ml ampicillin in a 250-ml culture flask (1:100 dilution). Grow at 30°C with vigorous aeration until absorbance at 550 nm reaches 0.4 to 0.6 OD/ml (~3.5 hr).

8. Remove a 1-ml aliquot of the culture (uninduced cells). Measure the optical density at 550 nm and harvest the cells by microcentrifuging 1 min at maximum speed, room temperature. Carefully remove all the spent medium with a pipet and store the cell pellet at –80°C.

9. Induce pL by adding 0.5 ml of 10 mg/ml tryptophan (100 µg/ml final) to remaining cells immediately.

10. Incubate 4 hr at 37°C. At hourly intervals during this incubation, remove 1-ml aliquots of the culture and harvest cells as in step 8.

11. Harvest the remaining cells from the culture 4 hr post-induction by centrifuging 10 min at 3000 rpm (e.g., in a Beckman J6 rotor), 4°C. Store the cell pellet at –80°C.

    Procedures for further analysis of these cells are outlined in the support protocols.

Verify induction

12. Resuspend the pellets from the induction intervals (steps 8 and 10) in 200 µl of SDS-PAGE sample buffer/OD_{550} cells. Heat 5 min at 70°C to completely lyse the cells and denature the proteins. Run the equivalent of 0.15 OD_{550} cells per lane (30 µl) on an SDS-polyacrylamide gel (*UNIT 10.1*).

BU IAST/ CENTRAL UNIVERSITY LIBRARY

13. Stain the gel 1 hr with Coomassie brilliant blue (*UNIT 10.5*). Destain the gel and check for expression.

Most thioredoxin fusion proteins are produced at levels that vary from 5% to 20% of the total cell protein. The desired fusion protein should exhibit the following characteristics: it should run on the gel at the mobility expected for its molecular weight; it should be absent prior to induction; and it should gradually accumulate during induction, with maximum accumulation usually occurring 3 hr post-induction at 37°C.

SUPPORT PROTOCOL 1

E. COLI LYSIS USING A FRENCH PRESSURE CELL

A small, 3.5-ml French pressure cell can be used as a convenient way to lyse *E. coli* cells. The whole-cell lysate can be fractionated into soluble and insoluble fractions by microcentrifugation. Other lysis procedures may be used—for example, sonication (*UNIT 6.6*) or treatment with lysozyme-EDTA (*UNIT 6.5*). For use of the larger 40-ml French pressure cell, see *UNIT 6.2*.

Materials

Cell pellet from 4-hr post-induction culture (see Basic Protocol)
20 mM Tris·Cl, pH 8.0 (*APPENDIX 2E*), 4°C
Lysis buffer: 20 mM Tris·Cl (pH 8.0) with protease inhibitors (optional)—0.5 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM *p*-aminobenzamidine (PABA), and 5 mM EDTA
SDS-PAGE sample buffer (see recipe)
French press and 3.5-ml mini-cell (SLM Instruments), 4°C

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*)

Lyse the cells

1. Resuspend cell pellet from 4-hr post-induction culture in 20 mM Tris·Cl, pH 8.0, to a concentration of 5 OD_{550}/ml.

Protease inhibitors can be included in the resuspension if desired. Cells can also be resuspended at densities of 100 OD_{550}/ml or greater; however, at high densities cell lysis may be less efficient.

2. Place 1.5 ml resuspended cell pellet in the 3.5-ml French pressure cell. Hold the cell upside down with the base removed, the piston fully extended downwards, and the outlet valve handle that holds the nylon ball seal in the open position (loose).

Before filling the pressure cell, check that the nylon ball, which seals the outlet port and sits on the end of the outlet valve handle, is not deformed. If it is, replace it with a new one. Both the condition of the nylon ball and its seat in the pressure-cell body are critical for the success of the procedure.

3. Bring the liquid in the pressure cell to the level of the outlet port by raising the piston slowly to expel excess air from the cell. With the outlet valve open and at the same time maintaining the piston in position, install the pressure-cell base. Gently close the outlet valve.

CAUTION: *Do not over-tighten the valve as this will deform the nylon ball and may irreparably damage its seat on the pressure-cell body.*

4. Turn the sealed cell right-side-up and place it in the hydraulic press.

5. Turn the pressure regulator on the press fully counter-clockwise to reset it to zero pressure. Set the ratio selector to medium. Turn on the press.

CAUTION: *The larger (50-ml) pressure cell is usually used with the selector set on high. The small (3.5-ml) cell is only used on medium ratio.*

ECU IASI CENTRAL UNIVERSITY LIBRARY

6. Slowly turn the pressure regulator clockwise until the press just begins to move. Allow the press to compress the piston.

   The press will stop moving after a few seconds.

7. Position a collection tube under the pressure-cell outlet. Slowly increase the pressure in the cell by turning the pressure regulator clockwise. Monitor the reading on the gauge and increase the pressure to 1000 on the dial, corresponding to an internal cell pressure of 20,000 lb/in^2.

8. While continuously monitoring the gauge, very slowly open the outlet valve until lysate begins to trickle from the outlet.

   The lysate should flow slowly and smoothly, and the cell pressure should not drop more than 100 divisions on the dial.

   At 20,000 lb/in^2 and 5 OD_{550}/ml, cell lysis will be complete after one passage through the press. Lower pressures and/or higher cell densities may require a second passage.

Fractionate the lysate

9. Remove a 100-µl aliquot of the lysate and freeze at –80°C (whole-cell lysate).

10. Fractionate the remainder of the lysate by microcentrifuging 10 min at maximum speed, 4°C.

11. Remove a 100-µl aliquot of the supernatant and freeze at –80°C (soluble fraction). Discard the remainder of the supernatant.

    Because this is a pilot experiment, it would not produce enough material to warrant saving any remaining supernatant.

12. Resuspend the pellet in an equivalent volume of lysis buffer. Remove a 100-µl aliquot and freeze at –80°C (insoluble fraction).

13. Lyophilize the 100-µl aliquots to dryness in a Speedvac evaporator. Solubilize in 100 µl SDS-PAGE sample buffer. Analyze 30-µl samples by SDS-PAGE (*UNIT 10.1*).

    This crude fractionation provides a fairly reliable indication of whether a protein has folded correctly. Usually proteins in the soluble fraction have adopted a correct conformation and proteins in the insoluble fraction have not. However, occasionally proteins found in the soluble fraction are not truly soluble; instead they form aggregates that do not pellet in the microcentrifuge. Conversely, sometimes a protein found in the insoluble fraction may be there because it has an affinity for cell-wall components and cell membranes, and it may not be intrinsically insoluble. Occasionally proteins can be recovered from these insoluble fractions by extracting with agents such as mild detergents.

OSMOTIC RELEASE OF THIOREDOXIN FUSION PROTEINS

SUPPORT PROTOCOL 2

Thioredoxin and some thioredoxin fusion proteins can be released with good yield from the *E. coli* cytoplasm by a simple osmotic shock procedure.

Materials

Cell pellet from 4-hr post-induction cultures (see Basic Protocol)
20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA/20% (w/v) sucrose, ice-cold
20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA, ice-cold

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*)

1. Resuspend cell pellet from 4-hr post-induction cultures at a concentration of 5 OD_{550}/ml in ice-cold 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA/20% sucrose. Incubate 10 min on ice.

2. Microcentrifuge 30 sec at maximum speed, 4°C, to pellet the cells.

BUNGE/CENTRAL UNIVERSITY LIBRARY

3. Discard the supernatant and gently resuspend the cells in an equivalent volume of ice-cold 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA. Incubate 10 min on ice and mix occasionally by inverting the tube.

   Osmotic release from the cytoplasm occurs at this stage.

4. Microcentrifuge 30 sec at maximum speed, 4°C. Save the supernatant (osmotic shockate). Resuspend the cell pellet in an equivalent volume 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA (retentate).

5. Lyophilize 100-µl aliquots of osmotic shockate and retentate to dryness in a Speedvac evaporator.

6. Solubilize each in 100 µl SDS-PAGE sample buffer. Analyze 30-µl aliquots by SDS-PAGE (*UNIT 10.1*).

   The osmotic-shock procedure provides a substantial purification step for some thioredoxin fusion proteins. This procedure will remove most of the contaminating cytoplasmic proteins as well as almost all of the nucleic acids. However the shockate will contain as contaminants about half of the cellular elongation factor-Tu (EF-Tu) and most of the E. coli periplasmic proteins.

SUPPORT PROTOCOL 3

PURIFICATION OF THIOREDOXIN FUSION PROTEINS BY HEAT TREATMENT

Wild-type thioredoxin is resistant to prolonged incubations at 80°C. A subset of thioredoxin fusion proteins also exhibit corresponding thermal stability, and heat treatment at 80°C can sometimes be used as an initial purification step. Under these conditions the majority of contaminating *E. coli* proteins are denatured and precipitated.

Materials

Cell pellet from 4-hr post-induction cultures (See Basic Protocol)
20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA
SDS-PAGE sample buffer (see recipe)

80°C water bath
10-ml glass-walled tube

Additional reagents and equipment for lysis using a French pressure cell (Support Protocol 1) and SDS-PAGE (*UNIT 10.1*)

1. Resuspend cell pellet from 4-hr post-induction cultures at a concentration of 100 OD_{550}/ml in 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA.

   It is important to start off with a high protein concentration in the lysate to ensure efficient precipitation of denatured proteins.

2. Lyse the cells at 20,000 lb/in^2 in a French pressure cell (see Support Protocol 1, steps 2 to 8). Collect whole-cell lysate in a 10-ml glass-walled tube.

3. Incubate whole-cell lysate 10 min at 80°C. Remove 100-µl aliquots after 30 sec, 1 min, 2 min and 5 min and plunge immediately into ice. At 10 min, plunge the remaining heated lysate into ice.

   A glass-walled tube (not plastic) provides good thermal conductivity to provide a rapid rise in temperature to 80°C and then a rapid drop in temperature to 4°C. A suitable volume to use in a 10-ml glass tube is 1.5 ml lysate. For large-scale work, a glass-walled vessel should be used and the lysate should be mixed well during both heat treatment and cooling.

4. Microcentrifuge the aliquots 10 min at maximum speed, 4°C to pellet heat-denatured, precipitated proteins.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

5. Remove 2-µl aliquots of the supernatants and add 28 µl SDS-PAGE sample buffer. Analyze the samples by SDS-PAGE (*UNIT 10.1*) to determine the heat stability of the fusion protein and the minimum time of heat treatment required to obtain a good purification.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Casamino Acids (CAA), 2% (w/v)

20 g Casamino Acids (Difco-certified)
H_2O to 1 liter
Autoclave or filter sterilize through a 0.45-µm filter
Store ≤2 months at room temperature

Do not use technical-grade Casamino Acids because it has a higher NaCl content.

CAA/glycerol/ampicillin 100 medium

800 ml 2% (w/v) Casamino Acids (see recipe; 1.6% final)
100 ml 10× M9 salts (see recipe; 1× final)
100 ml 10% (v/v) glycerol (sterile; 1% final)
1 ml 1 M $MgCl_2$ (sterile; 1 mM final)
0.1 ml 1 M $CaCl_2$ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
10 ml 10 mg/ml ampicillin (sterile; 100 µg/ml final)
Prepare fresh

IMC medium

200 ml 2% (w/v) Casamino Acids (see recipe; 0.4% final)
100 ml 10× M9 salts (see recipe; 1× final)
40 ml 20% (w/v) glucose (sterile; 0.5% final)
1 ml 1 M $MgCl_2$ (sterile; 1 mM final)
0.1 ml 1 M $CaCl_2$ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
658 ml glass-distilled H_2O (sterile)
10 ml 10 mg/ml ampicillin (sterile; optional; 100 µg/ml final)
Use fresh

IMC plates

15 g agar [Difco; 1.5% (w/v)]
4 g Casamino Acids [Difco-certified; 0.4% (w/v)]
858 ml glass-distilled H_2O (sterile)
Autoclave 30 min
Cool in a 50°C water bath
100 ml 10× M9 salts (see recipe; 1× final)
40 ml 20% (w/v) glucose (sterile; 0.5% final)
1 ml 1 M $MgCl_2$ (sterile; 1 mM final)
0.1 ml 1 M $CaCl_2$ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
10 ml 10 mg/ml ampicillin (sterile; optional; 100 µg/ml final)
Mix well and pour into petri plates
Store ≤1 month at 4°C

B. C. U. „M. EMINESCU" IAȘI
CHIMIE

BCU IAȘI / CENTRAL UNIVERSITY LIBRARY

M9 salts, 10×

60 g Na_2HPO_4 (0.42 M)
30 g KH_2PO_4 (0.24 M)
5 g NaCl (0.09 M)
10 g NH_4Cl (0.19 M)
H_2O to 1 liter
Adjust pH to 7.4 with NaOH
Autoclave or filter sterilize through a 0.45-µm filter
Store ≤6 months at room temperature

SDS-PAGE sample buffer

15% (v/v) glycerol
0.125 M Tris·Cl, pH 6.8 (*APPENDIX 2E*)
5 mM Na_2EDTA
2% (w/v) SDS
0.1% (w/v) bromphenol blue
1% (v/v) 2-mercaptoethanol (2-ME; add immediately before use)
Store indefinitely at room temperature

Tryptophan, 10 mg/ml

Heat 500 ml glass-distilled H_2O to 80°C. Stir in 5 g L-tryptophan until dissolved. Filter sterilize the solution through a 0.45 µm filter and store ≤6 months in the dark at 4°C.

COMMENTARY

Background Information

Two significant problems plague researchers who hope to express heterologous proteins in *Escherichia coli*: inefficient initiation of translation of many eukaryotic mRNA sequences on bacterial ribosomes (Stormo et al., 1982), and proteins that often form insoluble aggregates, called inclusion bodies, that are composed of misfolded or denatured proteins (Mitraki and King, 1989). Although successful protocols for refolding eukaryotic proteins from inclusion bodies can be developed, the process is always uncertain and usually time-consuming; in most instances it is preferable to prevent inclusion-body formation in the first place. The use of *trxA* fusions provides a solution to both problems.

Inefficient initiation of translation of eukaryotic messages in *E. coli* can often be improved by modifying sequences at the 5′ end of the gene. A more reliable technique that avoids the problem entirely is to use a gene-fusion strategy in which the gene of interest is linked in-frame to the 3′ end of a highly translated partner gene. In this case protein synthesis always initiates on the same efficiently translated fusion-partner mRNA, thus ensuring high-level expression. Some earlier gene-fusion expression systems, for example the *trpE* and *lacZ* systems, offer very reliable ways of producing large quantities of any desired eukaryotic protein. However, these gene-fusion systems still suffer from the pervasive inclusion-body problem. They are thus mainly useful for the production of antigens, rather than correctly folded, biologically active proteins. More recently the maltose binding protein (MBP; Riggs, 1994; *UNIT 5.1*) and glutathione-*S*-transferase (GST) gene fusion expression systems (see *UNIT 6.6*) have proven more successful in producing soluble fusion proteins; these systems retain the translation advantage of the earlier fusion systems. Apart from the obvious advantages in making a correctly folded product, the synthesis of soluble fusion proteins also allows for the development of generic purification schemes based on some unique property of the fusion partner.

Why would any particular eukaryotic protein produced in the *E. coli* cytoplasm be more soluble when it is linked to a fusion partner than it would be by itself? It is likely that physical properties of the fusion-partner protein are important, with efficient self-folding and high solubility being useful in this role. It is possible that some good fusion partners (proteins that fold efficiently and are highly soluble), by virtue of their desirable physical qualities, are able to keep folding intermediates of linked heterologous proteins in solution long enough for

PCU IASRI CENTRAL UNIVERSITY LIBRARY

them to adopt their correct final conformations. In this respect the fusion partner may serve as a covalently joined chaperone protein, in many ways fulfilling the role of authentic chaperone proteins (McCoy, 1992), analogous to the covalent chaperone role proposed for the N-terminal pro regions of a number of protein precursors (Silen et al., 1989; Shinde et al., 1993).

Many of the known properties of *E. coli* thioredoxin (Holmgren, 1985) suggested that it would make a particularly effective fusion partner in an expression system. First, thioredoxin, when overproduced from plasmid vectors, can accumulate to 40% of the total cellular protein, yet even at these expression levels all of the protein remains soluble. Second, the molecule is small (11,675 M_r) and would contribute a relatively modest amount to the total mass of any fusion protein, in contrast to other systems such as the *lacZ* system. Third, the tertiary structure of thioredoxin (Katti et al., 1990) reveals that both the N- and C-termini of the molecule are accessible on the surface and in good position to link to other proteins. The structure also shows that the molecule has a very tight fold, with >90% of its primary sequence involved in strong elements of secondary structure. This provides an explanation for thioredoxin's observed high thermal stability (T_m = 85°C), and suggests that the molecule might possess the robust folding characteristics that could make it a good fusion-partner protein. In support of this view, complete thioredoxin domains are found in a number of naturally occurring multidomain proteins, including *E. coli* DsbA (Bardwell et al., 1991), the mammalian endoplasmic reticulum proteins ERp72 (Mazzarella et al., 1990), and protein disulfide isomerase (PDI; Edman et al., 1985). These proteins can all be considered as natural precedents for thioredoxin fusion proteins.

The synthesis of small peptides in *E. coli* is often difficult, with the products frequently being extensively degraded or insoluble. The thioredoxin tertiary structure revealed that the characteristic active site, —CGPC—, protrudes from the body of the protein as a surface loop, with few interactions with the rest of the molecule. The loop does not seem to contribute to the overall stability of thioredoxin, so the production of peptides as insertions at this site was an attractive possibility. In this location they would be protected from host-cell amino- and carboxypeptidases, and thioredoxin's high solubility should help keep them in solution. In addition, the conformation of peptides inserted at this position would be constrained, which could be an advantage for applications in which it is desirable for the peptide to adopt a particular form.

Thioredoxin has indeed proven to be an excellent partner for the production of soluble fusion proteins in the *E. coli* cytoplasm (LaValllie et al., 1993b). Figure 6.7.2 demonstrates the production of soluble fusion proteins between thioredoxin and eleven human and murine cytokines and growth factors using the *trxA* vectors. All of these mammalian proteins had been previously produced in *E. coli* only as insoluble inclusion bodies. As thioredoxin fusions, the growth factors are not only made in a soluble form, but in most cases they are also biologically active in in vitro assays.

Experience gained while working with these and a number of other *trxA* fusion proteins shows that two further characteristics of thioredoxin can be exploited as purification tools. The first is the inherent thermal stability of the molecule, a property that is retained by some thioredoxin fusion proteins. This enables heat treatment to be used as an effective purification step. The second additional property relates to thioredoxin's cellular location. Although *E. coli* thioredoxin is a cytoplasmic protein, it has been shown to occupy a special position within the cell—it is primarily located on the cytoplasmic face of the adhesion zones that exist between the inner and outer membranes of the *E. coli* cell envelope (Lunn and Pigiet, 1982). From this location thioredoxin is quantitatively released to the exterior of the cell by simple osmotic shock or freeze/thaw treatments, a remarkable property that is retained by some thioredoxin fusion proteins, thus providing a simple purification step.

A more generic method for purification of any soluble thioredoxin fusion employs a modified form of thioredoxin (called "His-patch Trx"), which has been designed to bind to metal chelate resins (Lu et al., 1996; *UNIT 9.4*).

If the fusion protein is soluble after lysis or osmotic release (see Support Protocols 1 and 2), then any of the conventional purification methods described in Chapter 8 can be used. Alternatively, an affinity-based purification method can be used where an immobilized arsenical compound forms an adduct with the redox-sensitive vicinal dithiols present at the active site of thioredoxin. This affinity matrix was originally prepared by covalently linking 4-aminophenylarsenine oxide to cyanogen bromide–activated Sepharose 4B (Hannestad et al., 1982). Hoffman and Lane (1992) coupled the same ligand to Affi-Gel 10 (Bio-Rad); the

BOSTON UNIVERSITY LIBRARY

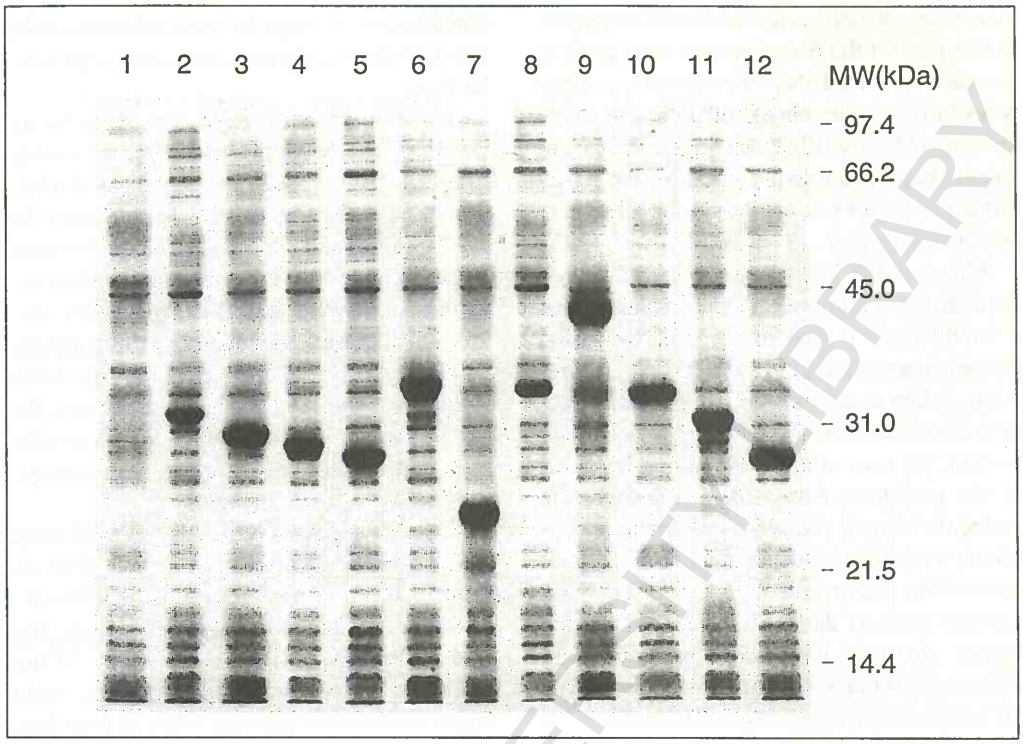


Figure 6.7.2 Expression of thioredoxin gene fusions. The gel shows proteins found in the soluble fractions derived from *E. coli* cells expressing eleven different thioredoxin gene fusions. Lane 1, host *E. coli* strain GI724 (negative control, 37°C); lane 2, murine interleukin-2 (IL-2; 15°C); lane 3, human IL-3 (15°C); lane 4, murine IL-4 (15°C); lane 5, murine IL-5 (15°C); lane 6, human IL-6 (25°C); lane 7, human MIP-1a (37°C); lane 8, human IL-11 (37°C); lane 9, human macrophage colony-stimulating factor (M-CSF; 37°C); lane 10, murine leukemia inhibitory factor (LIF; 25°C); lane 11, murine steel factor (SF; 37°C); and lane 12, human bone morphogenetic protein-2 (BMP-2; 25°C). Temperatures in parentheses are the production temperature chosen for expressing each fusion. This is a 10% SDS-polyacrylamide gel, stained with Coomassie brilliant blue.

resultant affinity matrix has a 10-atom spacer which makes it more useful for affinity chromatography. The resin is commercially available from Invitrogen (ThioBond) and Sigma. Protocols detailing the affinity-based purification approach can be obtained on-line from the Invitrogen web site (*http://www.invitrogen.com/manuals.html*), and a recent example from the literature is the expression of human glutamate decarboxylase (Papouchado et al., 1997). Other specific purification approaches include immunoaffinity chromatography using antibodies against either the target protein or against the thioredoxin moiety (monoclonal antibodies against thioredoxin are commercially available). Thioredoxin fusions with additional histidine tags or thioredoxin fusions based on mutant thioredoxins with metal-chelating affinity (Lu et al., 1996) can be purified using metal-chelate chromatography (see *UNIT 9.4*). Finally, if the fusion protein is insoluble, then attempts to solubilize and fold the fusion protein can be made prior to purification, in a technique analogous to that used with insoluble GST fusion proteins (see *UNIT 6.6*). Thioredoxin from *E. coli* is a very soluble protein that can be readily denatured and refolded (Kelley et al., 1987).

Critical Parameters

Lack of protein solubility leading to inclusion-body formation in *E. coli* is a complex phenomenon with many contributing factors: simple insolubility as a result of high-level expression, insolubility of protein-folding intermediates, lack of appropriate bacterial chaperone proteins, and lack of glycosylation mechanisms in the bacterial cytoplasm. Fusion of heterologous proteins to thioredoxin or to other fusion partners can help address most of these solubility issues. However, another important factor contributing to inclusion body formation is the inability to form essential disulfide bonds in the reducing environment of the bacterial cytoplasm, which leads to incorrect folding. Thermal lability of even correctly

ECU IAS CENTRAL UNIVERSITY LIBRARY

folded heterologous proteins in the absence of these stabilizing disulfide cross-links is a significant problem, so the expression of fusion genes should be attempted over a wide range of temperatures, even as low as 15°C (the limit for *E. coli* growth is ~8°C). Thermal denaturation is a time-dependent process, so it is also prudent to monitor the solubility of the expressed fusion protein over the time course of induction.

A great many proteins contain distinct structural domains. For example, hormone receptor proteins usually have an extracellular ligand-binding domain, a transmembrane region, and an intracellular effector domain. Sometimes, expressing these domains individually as fusion proteins can yield better results than expressing the entire protein. The exact positions chosen for boundaries of the domains to be expressed in the fusion protein are important and can be determined from a knowledge of the tertiary structure of the protein of interest, by homology comparisons with similar proteins, by limited proteolysis or other domain-mapping experiments, or empirically by generating multiple fusions that test different boundary positions.

It is important to be consistent in treating samples for loading on gels. For example, using different heating conditions from one experiment to the next can result in a mobility shift for the protein of interest.

Anticipated Results

Thioredoxin fusion protein yields are usually in the range of 5% to 20% of total cell protein. At these expression levels, a 1-liter induction culture in a shaker flask will yield ~3 g (wet weight) of cells, 300 mg total protein, and 15 to 60 mg of thioredoxin fusion protein. The final recovered yield will depend on factors such as solubility of the fusion protein and the efficiency of downstream purification procedures.

Time Considerations

From a single colony on a plate, the basic induction protocol requires an overnight growth to prepare a liquid inoculum and a 3.5-hr preinduction growth at 30°C the next day, followed by a 4-hr 37°C induction period. These times are significantly longer if lower induction temperatures are required (see Table 6.7.1). Lysis of a sample in the French pressure cell should require ≤ 5 min, and both the heat-treatment and osmotic-shock procedures require <1 hr each. SDS-PAGE takes 2.5 hr.

Literature Cited

Bardwell, J.C.A., McGovern, K., and Beckwith, J. 1991. Identification of a protein required for disulfide bond formation in vivo. *Cell* 67:581-589.

Edman, J.C., Ellis, L., Blacher, R.W., Roth, R.A., and Rutter, W.J. 1985. Sequence of protein disulphide isomerase and implications of its relationship to thioredoxin. *Nature* 317:267-270.

Hannestad, U., Lundqvist, P., and Sorbo, B. 1982. An agarose derivative containing an arsenical for affinity chromatography of thiol compounds. *Anal. Biochem.* 126:200-204.

Hoffman, R.D. and Lane, M.D. 1992. Isodophenylarsine oxide and arsenical affinity chromatography: New probes for dithiol proteins. *J. Biol. Chem.* 267:14005-14011.

Holmgren, A. 1985. Thioredoxin. *Ann. Rev. Biochem.* 54:237-271.

Katti, S.K., LeMaster, D.M., and Eklund, H. 1990. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 angstroms resolution. *J. Mol. Biol.* 212:167-184.

Kelley, R.F., Shalongo, M., Jagannadham, M.V., and Stellwagen, E. 1987. Equilibrium and kinetic measurements of the conformational transition of reduced thioredoxin. *Biochemistry* 26:1406-1411.

LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., and McCoy, J.M. 1993a. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268:23311-23317.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993b. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology* 11:187-193.

Lu, Z., DiBlasio-Smith, E.A., Grant, K.L., Warne, N.W., LaVallie, E.R., Collins-Racie, L.A., Follettie, M.T., Williamson, M.J., and McCoy, J.M. 1996. Histidine patch thioredoxins. Mutant forms of thioredoxin with metal chelating affinity that provide for convenient purifications of thioredoxin fusion proteins. *J. Biol. Chem.* 271:5059-5065.

Lunn, C.A. and Pigiet, V.P. 1982. Localization of thioredoxin from *Escherichia coli* in an osmotically sensitive compartment. *J. Biol. Chem.* 257:11424-11430.

Mazzarella, R.A., Srinivasan, M., Haugejorden, S.M., and Green, M. 1990. ERp72, an abundant luminal endoplasmic reticulum protein, contains three copies of the active site sequences of protein disulfide isomerase. *J. Biol. Chem.* 265:1094-1101.

McCoy, J.M. 1992. Heat-shock proteins and their potential uses for pharmaceutical protein production in microorganisms. *In* Stability of Protein Pharmaceuticals, Part B. (T. Ahern and M. Manning, eds.) pp 287-316. Plenum Press, New York.

BOSTON / CONTRACT UNIVERSITY LIBRARY

Mieschendahl, M., Petri, T., and Hanggi, U. 1986. A novel prophage independent trp regulated lambda pL expression system. *Bio/Technology* 4:802-808.

Mitraki, A. and King, J. 1989. Protein folding intermediates and inclusion body formation. *Bio/Technology* 7:690-697.

Norrander, J., Kempe, T., and Messing, J. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* 26:101-106.

Papouchado, M.L., Valdez, S.N., Ghiringhelli, D., Poskus, E., and Ermacora, M.R. 1997. Expression of properly folded human glutamate decarboxylase 65 as a fusion protein in Escherichia coli. *Eur. J. Biochem.* 246:350-359.

Riggs, P. 1994. Expression and purification of maltose-binding protein fusions. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.F. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp.16.6.1-16.6.14. John Wiley & Sons, New York.

Shatzman, A.R., Gross, M.S., and Rosenberg, M. 1990. Expression using vectors with phage λ regulatory sequences. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.F. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 16.3.1-16.3.11. John Wiley & Sons, New York.

Shimatake, H. and Rosenberg, M. 1981. Purified λ regulatory protein cII positively activates promoters for lysogenic development. *Nature* 292:128-132.

Shinde, U., Chatterjee, S., and Inouye, M. 1993. Folding pathway mediated by an intramolecular chaperone. *Proc. Natl. Acad. Sci. U.S.A.* 90:6924-6928.

Silen, J.L., Frank, D., Fujishige, A., Bone, R., and Agard, D.A. 1989. Analysis of prepro-α-lytic protease expression in *Escherichia coli* reveals that the pro region is required for activity. *J. Bacteriol.* 171:1320-1325.

Stormo, G.D., Schneider, T.D., and Gold, L. 1982. Characterization of translation initiation sites in *E. coli*. *Nucl. Acids Res.* 10:2971-2996.

Key Reference

Ausubel, F.M., Brent, R., Kingston, R.F., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) 1997. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

A source of protocols for cloning and analyzing DNA.

Internet Resources

http://www.invitrogen.com/manuals.html

Source for protocols on affinity-based purification.

Contributed by John McCoy
and Edward LaVallie
Genetics Institute
Cambridge, Massachusetts

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHARACTERIZATION OF RECOMBINANT PROTEINS

7

INTRODUCTION 7.0.1

7.1 Overview of the Characterization of Recombinant Proteins 7.1.1
Common Sources of Contamination 7.1.1
Chemical Identity 7.1.2
Physical and Conformational Characterization 7.1.3
Analysis of N-Terminal Heterogeneity of Recombinant Proteins Produced in *E. coli* 7.1.9
Literature Cited 7.1.11

7.2 Determining the Identity and Purity of Recombinant Proteins by UV Absorption Spectroscopy 7.2.1
Basic Protocol 1: Analysis of Proteins Using Near-UV Spectrophotometry 7.2.2
Support Protocol: Interpreting Protein Near-UV Spectra 7.2.5
Basic Protocol 2: Analysis of Proteins Using an HPLC with a Diode-Assay Detector 7.2.10
Reagents and Solutions 7.2.13
Commentary 7.2.13

7.3 Determining the Identity and Structure of Recombinant Proteins 7.3.1
Basic Protocol 1: Peptide Mapping by HPLC 7.3.2
Alternate Protocol: Peptide Mapping by High-Resolution SDS-PAGE 7.3.5
Support Protocol: Digestion of Protein in an SDS-Polyacrylamide Gel Slice 7.3.6
Basic Protocol 2: Determining the Disulfide Linkage Pattern of a Recombinant Protein 7.3.8
Basic Protocol 3: Determining Charge Heterogeneity by Two-Dimensional Titration Curve Analysis 7.3.13
Reagents and Solutions 7.3.15
Commentary 7.3.17

7.4 Transverse Urea-Gradient Gel Electrophoresis 7.4.1
Basic Protocol 7.4.1
Reagents and Solutions 7.4.5
Commentary 7.4.6

7.5 Analytical Ultracentrifugation 7.5.1
Planning Experiment 7.5.4
Anticipated Results 7.5.6
Time Considerations 7.5.8
Help for the Novice 7.5.8
Literature Cited 7.5.9

7.6 Determining the CD Spectrum of a Protein 7.6.1
Strategic Planning 7.6.7
Basic Protocol: Recording a CD Spectrum 7.6.10
Support Protocol 1: Interpretation of Near-UV CD Spectra 7.6.15
Support Protocol 2: Interpretation of Far-UV CD Spectra 7.6.18
Commentary 7.6.19

continued

BCU IAM/CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

7.7 Determining the Fluorescence Spectrum of a Protein **7.7.1**
Strategic Planning 7.7.1
Basic Protocol 1: Recording a Fluorescence Emission Spectrum 7.7.4
Basic Protocol 2: Determination of Fluorescence Quenching 7.7.8
Support Protocol: Basic Theory and Interpretation of Fluorescence Spectra 7.7.11
Commentary 7.7.17

7.8 Light Scattering **7.8.1**
Static Light Scattering Analysis of Protein Solutions 7.8.2
Dynamic Light Scattering Analysis of Protein Solutions 7.8.8
Literature Cited 7.8.13
Key References 7.8.14

7.9 Measuring Protein Thermostability by Differential Scanning Calorimetry **7.9.1**
Strategic Planning 7.9.1
Basic Protocol: Differential Scanning Calorimetry 7.9.2
Calibration of the DSC Instrument 7.9.4
Support Protocol 1: Electrical Calibration 7.9.4
Support Protocol 2: Temperature Calibration Using Hydrocarbon Capsules 7.9.4
Support Protocol 3: Temperature Calibration Using Lipid Suspensions 7.9.4
Support Protocol 4: Maintenance and Cleaning of DSC Cells 7.9.5
Commentary 7.9.6

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 7
Characterization of Recombinant Proteins

INTRODUCTION

Once a protein has been purified, it is necessary to characterize the protein preparation. The initial and most fundamental analyses performed are checks of purity and of biological or enzymatic activity. Indeed, to purify a protein in the first place a specific assay for the protein must be available, and limited purity determinations are usually made en route to the finished product by SDS-PAGE monitoring of the sample at different purification stages. The first unit of this chapter (*UNIT 7.1*) provides an overview of the approaches and methods used for more detailed characterization of recombinant proteins. It should be noted that none of the techniques or methods mentioned are used uniquely for recombinant proteins, and protocols dealing with some of the methods are included elsewhere in this book. However, the emphasis of the discussion in *UNIT 7.1* (and of the protocols which will appear in future updates) is on addressing the detection and determination of the most pertinent sources of contamination and heterogeneity commonly associated with proteins isolated from recombinant sources.

The first section of the overview describes common sources of contamination, both proteinaceous and nonproteinaceous. Because these contaminants are derived from the host cell, their specific characteristics will depend on the particular expression system used (also see Chapter 5). Once the purity of the protein has been established, the primary structure (amino acid sequence) is checked by standard protein-chemical methodologies. Throughout the overview, the use of HPLC and mass spectrometry is emphasized; these are important tools for checking the primary structure and especially for detecting both nonspecific and specific post-translational modifications.

Because many proteins expressed in *E. coli* form insoluble aggregates known as inclusion bodies, before purification they must be extracted under conditions that unfold the protein (see *UNITS 6.1, 6.3 & 6.5*). After folding and purification of the protein, there is always the possibility that varying amounts of misfolded conformers may be present. Misfolded protein can be considered as a contaminant. A number of spectroscopic and other methods used to study protein conformation and structure are tabulated in the overview. Protocols describing the most commonly used of these methods are included in this chapter (see *UNITS 7.2, 7.4, 7.6, 7.7 & 7.9*).

Proteins expressed in *E. coli* can retain N-terminal methionine derived from the initiating *N*-formylmethionine. The retention or nonretention of N-terminal methionine is a potential source of heterogeneity, and is discussed in the last section of the overview (also see *UNIT 6.1*).

In *UNIT 7.2* the analysis of proteins using ultraviolet (UV) absorption spectroscopy is described. The first part deals with direct (or batch-mode) spectroscopic measurement, in which the purified protein is simply placed in a quartz cuvette and analyzed. Data collection and interpretation protocols allow assessment of the identity and, to a lesser degree, the purity of the protein. The second part of *UNIT 7.2* describes the analysis of proteins simultaneously separated by HPLC chromatography. This can be considered to be a two-dimensional method in which the first dimension is chromatography; the peaks thus separated are spectroscopically analyzed using diode array detection in the second

Contributed by Paul T. Wingfield and Roger H. Pain
Current Protocols in Protein Science (1998) 7.0.1-7.0.4
Copyright © 1998 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

dimension. Both batch-mode and column-based spectroscopic analysis complement purity analysis using SDS-PAGE (Chapter 10) in that the former methods can detect the presence of nonproteinaceous contaminants, especially nucleic acid.

The primary sequence of a recombinant protein is known from the cDNA coding sequence. However, the sequence of an isolated recombinant-derived protein may differ from the predicted sequence due to, for example, post-translational modifications; it is therefore important to confirm its primary sequence. The amount of analysis required to do this will depend on the intended use of the protein. For many investigators the analysis may consist of limited N-terminal and C-terminal sequencing, peptide mapping, and either amino acid analysis (*UNIT 3.2*) or molecular weight determination using mass spectrometry (Chapter 16). On the other hand, a full structural determination is usually required for proteins intended for pharmaceutical use.

UNIT 7.3 describes some techniques that give information on the primary structure. The method most commonly used to review the primary sequence of a protein is peptide mapping. Peptides derived from protein chemically or enzymatically fragmented are separated by HPLC or electrophoresis. The interpretation of the resultant chromatograms, or band patterns, is based on the specificity of the cleavage method (usually very selective) and the expected sequence of the protein. Often comparisons are made with well-characterized standards. The most common post-translation modification is the oxidation of paired sulfhydryl residues to form a disulfide bond. The technique of peptide mapping is usually used to determine which sulfhydryl residues are paired (the linkage pattern). Specific details of HPLC methodology (Chapter 11) and gel electrophoresis (Chapter 10) are detailed elsewhere and peptide analysis using mass spectrometry will be described in Chapter 16. The electrophoretic technique of two-dimensional titration curve analysis is also included in *UNIT 7.3*, as it can give information on the charged residue composition of the protein in addition to guiding the selection of ion-exchange resins used for protein purification.

The topic of protein conformation is addressed in *UNIT 7.4*, in which the electrophoretic method known as transverse urea-gradient gel electrophoresis (UGGE) is described. The unfolding of a protein exposed to a linear gradient of urea (usually 0 to 8 M) is monitored simply by observation of the electrophoretic mobility. Gel patterns can be analyzed to estimate conformational stability and help identify covalent or conformational heterogeneity. Spectroscopic methods such as circular dichroism (*UNIT 7.6*) and fluorescence (*UNIT 7.7*), and to a lesser degree, UV-spectroscopy (*UNIT 7.2*), are among the standard methods used to analyze protein conformation, and descriptions of these methods are planned for this chapter. However, for investigators not wishing or not able to invest in expensive and dedicated equipment, the gel electrophoretic approach described herein may be a useful alternative, although not as flexible or rigorous (potentially) as spectroscopic and calorimeter-based methods.

For a purified protein, determination of the molecular weight and physical homogeneity in solution are of fundamental importance. With recombinant proteins the molecular weight can be predicted from the cDNA coding sequence and rapidly confirmed by mass spectroscopy (Chapter 16). Determination of the molecular weight of the native protein in solution is not so straightforward, but is still fairly routine using the modern analytical ultracentrifuge. *UNIT 7.5* is an overview of analytical ultracentrifugation that includes a discussion of methods for determining both the molecular weight and hydrodynamic properties of proteins.

Light-scattering methods can also provide information about the native molecular weight, oligomeric composition, and overall conformation. Techniques using size-exclusion

BCU IAS / CENTRAL UNIVERSITY LIBRARY

chromatography with on-line light scattering detectors are especially useful for characterizing proteins and glycoproteins (for review see Wen et al., 1996). *UNIT 7.8* provides an overview of the various light-scattering methods used to characterize proteins and macromolecular complexes. It should be noted that this method, like analytical centrifugation, provides direct molecular weight information without the use of calibration standards or models.

Circular dichroism and fluorescence spectroscopy are popular methods for determining protein conformation. Basic protocols for determining the circular dichroic spectra of proteins are described in *UNIT 7.6* and for determining the fluorescence spectra in *UNIT 7.7*. Circular dichroism and fluorescence spectroscopy are similar to most other spectroscopic methods in that careful interpretation of the data is required together with an understanding of the limitations and pitfalls of the method. These points are emphasized throughout these units.

Differential scanning calorimetry (DSC), described in *UNIT 7.9*, is an instrumental method used to measure heat energy uptake (heat capacity) in proteins undergoing thermal denaturation. This is accomplished by varying the temperature of a dissolved protein sample and measuring the difference in heat generated by the protein compared to a reference cell that contains solvent alone. This heat difference is related to the conformational energy of the protein molecule. In practice, when proteins undergo thermal denaturation, typically at 40° to 80°C, there is usually heat generated (i.e., it is an exothermic process); this heat can be very accurately measured by modern DSC instruments. The unit discusses the theoretical background of the method and details methods for actually performing DSC. For background on basic thermodynamic principles consult, for example, Tinco et al. (1995), and for overview of DSC applications in biotechnology see Chowdhry and Cole (1989). The Internet sites of the various instrument manufacturers are also useful sources of information.

CONFORMATIONAL ANALYSIS OF RECOMBINANT-PRODUCED PROTEINS

The assays for many protein functions, for example in the case of cytokines, can be time consuming and of relatively low precision. As proteins depend for their function on the correct 3D fold of the polypeptide chain, demonstration that the conformation of the recombinant protein is the same as that of the authentic protein should provide a more convenient means of characterizing the product. The methods described in *UNITS 7.4, 7.6 & 7.7*, although providing average structural information only, as compared with X-ray crystallography or NMR, are sensitive indicators of conformation and conformational change and, as such, can provide a highly specific "fingerprint" of a protein. Such conformational analysis will follow assurance of chemical authenticity and homogeneity, as described in *UNITS 7.1-7.3* (and elsewhere in the book). Urea gradient gel electrophoresis (*UNIT 7.4*) is a hydrodynamic technique that monitors the degree of compactness of the protein. It allows unfolding and refolding to be studied, leading to characterization of the protein in terms of its stability and kinetics. Near-UV circular dichroism (CD) results in spectra that are characteristic of the polar asymmetry of the environments of the aromatic residues, mainly tryptophan and tyrosine. Far-UV CD monitors the polypeptide backbone conformation, and hence, the secondary structure. Fluorescence spectra are determined chiefly by the polarity of the environment of the tryptophan and tyrosine residues and their interactions.The four techniques thus provide different, but complementary, fingerprints of the protein conformation that provide a high degree of assurance of nativeness within a few hours. For quality assurance of the native conformation and for general

BOJANA CENTRAL UNIVERSITY LIBRARY

characterization, it is important to employ as broad a range of the available techniques as possible.

Besides confirming the nativeness of the folded conformation of a sample of a protein that has already been characterized, either as the previously produced recombinant protein or as the authentic protein, the techniques described here offer a rapid means of obtaining structural and dynamic information, together with estimates of the stability of the protein.

LITERATURE CITED

Chowdhry, B.Z. and Cole, S.C. 1989. Differential scanning calorimetry: Applications in biotechnology. *Trends. Biotechnol.* 7:11-18.

Tinco, I., Sauer, K., and Wang, J.C. 1995. Physical Chemistry: Principles and Applications in Biological Sciences (3rd ed.). Prentice Hall, Upper Saddle River, N.J.

Wen, J., Arakawa, T. and Philo, J.S. 1996. Size-exclusion chromatography with on-line light scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal. Biochem.* 240:155-166.

Paul T. Wingfield and Roger H. Pain

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of the Characterization of Recombinant Proteins

With the advent of recombinant DNA technology and the availability of numerous expression systems, scientists are now able to prepare large amounts of proteins in relatively short periods of time. Many of the proteins produced by this technology are new and previously uncharacterized. Current developments in protein purification techniques, especially the use of affinity matrices, can eliminate a number of less-specific purification steps, accelerating this process. It is still challenging, however, to purify a protein to complete homogeneity, removing all proteinaceous and nonproteinaceous contaminants.

How clean a protein product needs to be largely depends on its final use. Some applications require exhaustive testing to ensure that no foreign material accompanies the protein. This is extremely important, for example, when the protein is to be used therapeutically, because even a minor contaminant can lead to side reactions that may be fatal. For such applications, it is therefore critical to evaluate the recombinant protein not only for its physical impurities, but also for its potential to cause immunogenic or allergic responses (Bristow, 1989). Strict guidelines for validating the purity of proteins for therapeutic use are in place at the U.S. Food and Drug Administration (Geisow, 1991; Garnick et al., 1991).

Such careful characterization may not be required for proteins used in research, where a slight impurity may not affect the activity of a protein. It is known, however, that failure of crystallization is often due to microheterogeneity of the protein. The most troublesome impurities result from deamidation of Asn and Gln, Cys oxidation, proteolysis, sequence polymorphisms, and variability in post-translational modifications (Wood, 1989).

A number of different expression systems are available for producing recombinant proteins, including bacteria (most commonly *E. coli*), yeast, insect cells, and mammalian cells infected with retroviruses (Chapter 5). Bacterial expression systems are generally the easiest to handle, and the first systems tried. Some proteins, however, require specific post-translational modifications (e.g., phosphorylation, glycosylation, and/or fatty acid acylation) for optimal activity, thus forcing the use of the other (eukaryotic) expression systems. Each expression system has its own set of advantages and disadvantages (Chapter 5), and each has its own set of likely contaminants. A comprehensive set of analytical techniques is available to ascertain the purity of the protein products.

COMMON SOURCES OF CONTAMINATION

Contaminating Proteins

It is critical to check that the sample is free of other contaminating proteins. The hardest to detect and eliminate are those of similar molecular weight, pI, and hydrophobicity. The best way to analyze for such impurities is by HPLC, using at least two complementary chromatographic separations (e.g., reversed-phase chromatography on a C4 column and ion-exchange chromatography, or the use of two types of ion-pairing agents in reversed phase) with the aid of a diode-array detector. Another powerful method is the use of high-resolution peptide mapping by HPLC, alone or in combination with electrospray mass spectrometry (Scoble and Martin, 1990; Carr et al., 1991) or matrix-assisted laser-desorption ionization time-of-flight (MALDI TOF) mass spectrometry (see *UNIT 11.6*).

If authentic pure protein is available for comparison, proteolytic digestion of both the authentic and the recombinant protein under the same conditions can be used to assess purity of the recombinant protein. Both proteins should generate the same fragmentation pattern; differences may be interpreted as being due to contaminants or as indicating chemical heterogeneity. Electrospray and MALDI TOF mass spectrometry eliminate the requirement for pure authentic protein, because the mass assignments for each proteolytic fragment of the recombinant protein are accurate enough on their own to indicate if the fragments are all derived from the parent protein. Any fragments that do not fit may belong to a contaminant or indicate incorrect primary sequence.

Proteolytically Clipped Proteins

Many proteins become proteolytically cleaved during the purification process, even in the presence of protease inhibitor cocktails, due to the ubiquitous distribution of proteolytic agents in cell lysates. In some circumstances it may be difficult to separate the fragments from the product of interest, especially when the protein is nibbled from one end or the other by

Copyright © 1995 by John Wiley & Sons, Inc.

BOKU-UNIVERSITÄT LIBRARY

exoproteases, as these products will be almost the same size and have nearly identical physical properties. This task becomes especially difficult when the level of contamination is below normal detection levels. For proteins used therapeutically, this problem may have major ramifications, because it is well known that degraded proteins, or those that are not folded correctly, may be more antigenic than their intact counterparts, causing the patient to produce antibodies to the protein itself (Walsh and Headon, 1994). Determination of the N- and C-terminal sequences may be required to ensure purity from this type of contamination.

Nucleic Acid

Host DNA may accompany the purified protein. The main concern with this contaminant is the possibility that an oncogene from the host may be present in the sample. DNA is generally efficiently removed from samples by chromatography through anion exchangers such as DEAE. Spectroscopic measurements at 260 nm will indicate the presence of nucleic acids.

Pyrogens

Bacterial endotoxins and lipopolysaccharides can often accompany a protein through its purification (Bristow, 1989). These agents, if present even in small amounts, can produce fever in patients. Thus, for proteins used therapeutically, it is important to prove that such contaminants are not present in the sample. Apart from chromatography by HPLC, sensitive immunological methods also exist to test for the most common of these contaminants. For example, bacterial proteins can be detected by using antibodies against total *E. coli* cell lysates in immunoblots of the recombinant protein of interest. If the presence of lipopolysaccharides is suspected, they usually can be removed by gel filtration or by using membranes of the appropriate molecular weight cutoff, as these contaminants are very hydrophobic and tend to aggregate.

For further details on the detection and analysis of nonproteinaceous contaminants, see Walsh and Headon (1994).

CHEMICAL IDENTITY

Checking Primary Structure of Recombinant Proteins

Once the purity of the recombinant protein has been ascertained, it is imperative to prove that the isolated protein indeed has the correct amino acid sequence. There are a number of analytical techniques that can be used to characterize the protein, each providing different but overlapping answers. Some of the techniques are more exacting than others. They all, however, generally require expensive instrumentation, such as HPLC machines or mass spectrometers, that may not be available in all laboratories. Other techniques are less precise in their measurements, but provide reasonable data if authentic protein is available as a control.

The first test is to determine the mass of the recombinant protein, especially as it compares to the authentic protein. This can be accomplished in most laboratories by SDS-PAGE. Although the mass determined in this manner is not very accurate, it can be compared to the control to determine whether the protein is close to the right size. A more precise measure of the mass can be obtained by mass spectrometry, either electrospray or MALDI TOF. These measurements are generally within 0.01% to 0.05% of the correct mass.

Next, it is useful to determine the isoelectric point (pI) of the recombinant protein. This can be accomplished by isoelectric focusing (IEF) gel electrophoresis (*UNIT 10.2*) using flat-bed agarose or acrylamide gels or slab gels. The pI is determined from a standard curve constructed using IEF standards (proteins of known pI). It is good practice to include the native protein, if available, in an adjacent lane. Differences in pI from the expected value are diagnostic for differences in charged amino acids.

A small sample of the protein (1 to 5 μg) should be analyzed for its amino acid composition. This can usually be done in a protein chemistry core facility, available at most universities, for a small fee. This analysis should provide confirmation that all the expected amino acids are present in the protein. A second sample (25 to 50 pmol) should be submitted for N-terminal amino acid sequencing. This sample can be provided on a PVDF membrane (electroblotted from an SDS gel), or in solution (provided the sample is free of salts). This analysis should verify that the protein starts with the expected sequence, an important consideration if the N-terminal Met is cleaved after translation. A blocked N-terminus, such as occurs when the terminal amino acid is post-translationally modified, requires special procedures to be identified (LeGendre et al., 1993).

High-resolution peptide maps (fingerprinting) provide information on internal sequences

by comparing digests of the recombinant and the authentic protein, both treated under the same conditions. The best diagnostic fingerprints are analyzed by HPLC, where the retention times of over a hundred different fragments can often be compared. A difference in amino acid sequence between the recombinant protein and the authentic protein is highly likely to result in a change in the mobility of at least one fragment. If an HPLC unit is not available, one can get relatively good information from analysis of digested proteins by high-resolution SDS-PAGE. In that case one should use multiple partial digests to compare fragments that are ≥2 kDa in size. The Tris-tricine gel system (*UNIT 10.1*) is recommended for this type of analysis.

New, powerful methods are being developed that combine peptide mapping with mass spectrometry (Pappin et al., 1993). In these procedures, the mass of each fragment is obtained with high accuracy, allowing direct comparison with the predicted sequence from the cDNA. Thus, it is not necessary to include the authentic protein in the analysis, unless post-translational modifications are expected.

Characterizing Post-Translational Modifications

It is important to determine the nature and location of all post-translational modifications of recombinant proteins. Among the most problematic are unwanted, nonspecific modifications, including oxidation of Met residues and chance deamidation of Asn or Gln. These modifications can easily be detected as retention time shifts when HPLC profiles of peptide maps of the recombinant protein are compared to those of the authentic protein. The presence of reducing agents, such as 2-mercaptoethanol or dithiothreitol, in the isolation buffer often prevents oxidation of Met.

There are situations in which one wants to determine if the correct post-translational modification has been made in the recombinant protein, especially if the modification is important for the activity of the protein. Usually, the expression system has been carefully selected to provide the appropriate modification. Among the most frequent post-translational modifications are phosphorylation, N-methylation, acetylation, glycosylation, farnesylation, and sialic acid capping of oligosaccharides.

Pinpointing these modifications can be time-consuming, requiring specific methods for each modification if traditional approaches are to be used (see Chapters 12 and 13 for protocols for determining sites of glycosylation and phosphorylation). The usual method is to digest the recombinant protein, separate the fragments by HPLC, then pick fragments suspected of being modified for further analysis by N-terminal sequencing using Edman chemistry. A major disadvantage to this approach is that some of the modifications are not stable when exposed to the chemicals used in sequencing, and will therefore be missed. For example, phosphorylated Ser and Thr residues are not easily detected because they undergo β-elimination of the phosphoryl group (Mercier et al., 1971). Instead, a system using radioactive tracers has been developed in which the assignment of the phosphorylation sites is based on the release of [^{32}P]Pi during Edman degradation (Wettenhall and Cohen, 1982).

Newer approaches are being developed that rely on mass spectrometry to pinpoint the location of the modifications (Wang et al., 1994). These approaches offer direct procedures for identifying the modified residues and should cut the time required for analysis.

PHYSICAL AND CONFORMATIONAL CHARACTERIZATION

Once the chemical identity of a protein has been established, its physical and conformational properties are probed using a variety of biophysical techniques. Unlike chemical analysis, most of these techniques require that the protein be pure, as otherwise the interpretation of the data can be difficult or impossible. Some of the basic questions that are addressed are: (1) Is the protein monodisperse? (2) What is the quaternary (subunit) structure? (3) Is the protein folded? (4) What are the secondary and tertiary structures of the protein?

The commonly used methods for studying protein conformation and structure are listed in Table 7.1.1 (nonspectroscopic) and Table 7.1.2 (spectroscopic). References have been included that emphasize the practical aspects of the various techniques. Because of the complexity of the subject matter, several outstanding textbooks that cover most of the theoretical background have also been cited (see Key References).

Physical Appearance and Absorbance

It is always useful to visually inspect purified protein placed in a glass or quartz cuvette; it may be obvious on holding the cuvette up to the light that the protein is colored, cloudy, or perhaps slightly opalescent. For recombinant

BOUNTIFUL ETRAL UNIVERSITY LIBRARY

Table 7.1.1 Nonspectroscopic Methods for Studying Protein Structure and Conformation

Method	Information obtained	Notes	Reference(s)
Urea gradient electrophoresis	Size, shape, and unfolding transitions	Amount of material required depends on size of gel and detection system. If antibodies are used, crude samples can be analyzed.	Goldenberg (1989)
Gel filtration	Size, shape, and molecular interactions	Most commonly used method for determining molecular radii. M_r can be determined if sedimentation coefficient is known. Conformation changes can be followed.	Ackers (1970); Ptitsyn et al. (1990)
Ultracentrifugation		Various sedimentation velocity measurements used to gain information on shape/conformation.	
Sedimentation velocity	Shape	Useful check of homogeneity. Determination of sedimentation coefficient (S). S is sensitive to protein conformation. Small differences in S between samples may be difficult to detect.	Van Holde (1975); Harding (1994); Ralston (1993)
Difference sedimentation	Shape	The measurement of very small differences in sedimentation coefficient (~0.01 S). Sensitive to very small conformation differences between samples.	Van Holde (1975)
Sedimentation equilibrium	Molecular weight, quaternary structure, and molecular interactions	The method of choice for direct determination of the molecular weight (precision ±3%). Also widely used for studying reversible self-association of proteins and ligand binding.	Van Holde (1975); Harding (1994); Ralston (1993); McRorie and Voelker (1993)
Accessibility assays			
Protease sensitivity	Conformation	Protease action on folded proteins depends on both the cleavage-site specificity and its accessibility. In unfolded proteins only the former is important.	Price and Johnson (1989)
Chemical reactivity	Conformation	Similar in principal to above except using chemical reagents specific for various amino acid side chains (e.g., sulfhydryl reagents).	Ballery et al. (1993)
^2H and ^3H-exchange	Conformation/dynamics	HPLC and NMR used to measure exchange reactions.	Englander and Kallenbach (1984)
Calorimetry			
Differential scanning	Conformation; thermal stability	Measurement of heat uptake (heat capacity) during change in temperature. Used for study of thermal unfolding of proteins and nucleic acids. In modern instruments sample requirements are modest: 1-2 ml at 0.5-2.0 mg/ml per scan. Proteins often irreversibly denature at neutral pH; this is sometimes circumvented by analysis at acid pH or by addition of low concentrations of chaotropes.	Sturtevant (1987); Cooper and Johnson (1994)
Isothermal titration	Molecular interactions	Direct measurement of the energetics of bimolecular interactions (including binding constants). Larger samples needed than for differential scanning.	Cooper and Johnson (1994)

continued

DCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 7.1.1 Nonspectroscopic Methods for Studying Protein Structure and Conformation, continued

Bimolecular interaction analysis	Molecular interactions	Bimolecular interaction analysis using the Pharmacia BIAcore instrument. Uses surface plasmon resonance to detect bimolecular interactions. Examples: protein ligand binding to receptor, protein binding to nucleic acid, and antibody-antigen interactions. Wet chemistry required to immobilize one of the reaction partners on a matrix.	Chaiken et al (1993); Cunningham and Wells (1993)
Immunological assays	Conformation	Affinity of an antibody for its antigen depends on the structural complementarity between the molecules. Changes in the antigen conformation result in changes in the binding affinity for appropriate antibodies. Monoclonal antibodies are usually used.	Figuet et al. (1989)

proteins isolated from *E. coli*, all of these observations are usually bad news. Running an absorption spectrum from 240 nm (in the near-UV) to 600 nm (in the visible region) will enable one to decide if the protein is partially aggregated or is contaminated with colored chromophores. The absorbance of a ~1 mg/ml protein solution (using a 1-cm path length cell) in the region of 400 and 320 nm should be <0.05; this is a spectral region where pure protein does not absorb, and high absorbance (>0.1) is usually due to light scattering from aggregated protein. It should be noted that light scattering from macroscopic particles increases with decreasing wavelength; hence, the detection of scattering artifacts is aided by measuring an absorption spectrum rather than single wavelengths. Filtration with a 0.2-μm-pore-size filter, or low-speed centrifugation, often removes the material responsible for light scattering. A clearer picture of what proportion of the sample is aggregated requires further analysis by gel-filtration (see discussion of Molecular Size and Subunit Structure) or analytical ultracentrifugation. The main absorption peak should be at ~280 nm and the A_{280}/A_{260} ratio should be ~1.5 to 2.0. If there is high absorbance at 260 nm, it may be due to nucleic acid or other chemical contaminants.

Molecular Size and Subunit Structure

The next step is to apply the sample to a gel-filtration column to see if the protein elutes in a single symmetrical peak as expected for a monodisperse protein (*UNIT 8.3*). The appearance of material eluting at the void volume of the column indicates highly aggregated protein. If the column is calibrated with molecular size markers, an approximate idea of the size can be obtained. It should be remembered that to get an accurate molecular weight, it is necessary to combine the gel-filtration data with a knowledge of the sedimentation coefficient, which must be obtained either by analytical ultracentrifugation or density gradient analysis (*UNIT 4.2*).

If the molecular weight of the native protein is known, the subunit composition can be determined, as the monomeric mass is known from the amino acid sequence translated from the DNA coding sequence. The only commonly used methods capable of directly determining molecular weight are sedimentation equilibrium and light scattering (osmotic pressure measurement also gives this information but is rarely used nowadays).

Determination of the molecular weight of recombinant glycoproteins by gel filtration is especially unreliable. If the amount of carbohydrate associated with the protein has been determined (e.g., by mass spectrometry), the molecular weight can be directly determined by sedimentation equilibrium. With ultracentrifugation, the association state of a glycoprotein (e.g., whether it is a monomer or dimer) can be determined even when the carbohydrate composition is unknown (Pennica et al., 1992; Shire, 1992).

Apart from some of the methods referred to above, a more modest approach, equipment-wise, to molecular weight determination is the use of native gel electrophoresis (*UNIT 10.3*).

Conformation and Structure

It is important to assess if the purified recombinant protein has been folded correctly,

EDINBURGH CENTRAL UNIVERSITY LIBRARY

Table 7.1.2 Spectroscopic Methods for Studying Protein Structure and Conformation

Method	Information obtained	Notes	Reference(s)
Ultraviolet absorption			
Standard spectroscopy	Secondary and tertiary structure	Method available to most laboratories. Modern instruments have diode array detectors and are computer controlled.	Wetlaufer (1962)
Difference spectroscopy	Conformational change	Spectral differences between perturbed (e.g., using pH, temperature, and denaturants) and nonperturbed samples give information on relative conformation.	Donovan (1969, 1973)
Fluorescence		Very sensitive; small samples needed compared to other spectroscopic methods.	
Steady-state	Secondary and tertiary structure	Intrinsic fluorescence mainly due to tryptophan. Proteins can be labeled with fluorophores (extrinsic fluorescence). Steady-state quenching studies give information on exposure of chromophores.	Chen et al. (1969); Haugland (1994)
Time-resolved	Conformational change	Trp fluorescence lifetimes (decay time) lie in the nanosecond region. Molecular dynamics can be studied by time-resolved quenching and depolarization.	Eftink and Ghiron (1981); Varley (1994)
Circular dichroism	Conformational change	One of the most commonly used techniques to compare overall (or gross) conformational properties of proteins.	
Far-UV	Secondary structure	Measurements in the range 240-180 nm. α-helix most accurately determined. Computer-aided analysis of spectra by various methods to estimate secondary structure contents.	Yang et al. (1986); Johnson (1990)
Near-UV	Tertiary structure	Measurements in the range 340-240 nm. Depends on environment of aromatic residues. Conformational fingerprint. Interpretation not easy, but useful for comparative purposes.	Strickland (1974)
Raman scattering	Secondary structure	Raman scattering (and infrared absorption) known as vibrational spectroscopy. Useful for large macromolecular assemblies, precipitates, and crystals. Secondary structure estimates using Raman Amide I and III spectra.	Bussian and Sander (1989); Williams (1986)
Infrared absorption	Secondary structure	Modern computerized Fourier transform (FT-IR) instrumentation used; allow measurements in H_2O (particulate samples can also be measured). Secondary structure from Amide I spectra. Better than circular dichroism for β-structure estimation.	Haris and Chapman (1992, 1994); Surewicz et al. (1993)
Light scattering		Direct M_r determination (cf. sedimentation equilibrium). Especially useful for large macromolecular assemblies.	

continued

ECU IASI / CENTRAL UNIVERSITY LIBRARY

Table 7.1.2 Spectroscopic Methods for Studying Protein Structure and Conformation, continued

Static light scattering	Molecular weight and shape	Classical light scattering. Modern instruments use lasers. Samples must be dust-free. Determination of M_r requires that the refractive index increment (dn/dc) be known. M_r distributions and conformational information can be obtained by combining gel filtration with an in-line multi-angle light scattering detector.	Harding (1994)
Dynamic quasi-elastic	Shape and molecular weight	Also known as photon correlation spectroscopy. The translation diffusion coefficient (D) is measured. Knowledge of the sedimentation coefficient (S) is required to determine M_r.	Harding (1994); Phillies (1990)
Nuclear magnetic resonance			
^1H	Secondary structure; 3-D structure of small proteins and polypeptides (<10 kDa)	Large sample requirements. Proteins must be very pure and stable.	Clore and Gronenborn (1992)
Multidimensional	Secondary structure; 3-D structure of proteins 10-30 kDa	Proteins uniformly labeled with combinations of ^{15}N, ^{13}C, and ^2H by expression in *E. coli* using minimal medium (see *UNIT 5.3*). 2-D ^1H-^{15}N correlation spectra are very useful as protein conformational fingerprints.	Clore and Gronenborn (1994)
X-Ray diffraction	3-D structure at atomic resolution	Proteins must be crystallized. Commercially available Crystal screen kit widely used.	McRee (1993); Ducruix and Giege (1992); Hampton Research
Scanning transmission electron microscopy	Molecular weight	Scanning transmission electron microscopy; mass determination of individual molecules. Useful for large particles beyond the mass range of sedimentation equilibrium.	Thomas et al. (1994)

especially if it was prepared from inclusion bodies (insoluble aggregates of the over-expressed protein), which are often found in bacterial expression systems (*UNITS 6.3-6.5*). Spectroscopic methods and calorimetry are the usual techniques used to monitor or measure protein conformation. Urea-gradient electrophoresis is also used and has the advantage that pure protein is not required. It is advantageous—as in mapping the chemical identity of the protein—to have some type of standard available for comparative purposes. The power of the spectroscopic methods, especially circular dichroism (CD) and fluorescence, lies in their ability to detect differences rather than to make absolute determinations. The standard may be the natural or authentic protein or may be the same recombinant protein separately prepared and analyzed by independent methodologies.

Apart from the use of protein standards, model compounds are often used as conformation standards (especially for fluorescence and UV spectroscopy). These may be the free amino acids tryptophan and tyrosine: the model protein is constructed using the appropriate molar ratio of tryptophan and tyrosine and represents a fully denatured protein in which the aromatic residues are fully exposed to the solvent. Alternatively, the recombinant protein can be denatured with 4 to 8 M guanidine·HCl (see *APPENDIX 3A*). Comparison of either simulated or denatured protein with the test protein will give some indication of whether the protein is folded.

The most commonly used spectroscopic methods are CD and fluorescence. Nuclear magnetic resonance (NMR) spectroscopy measurements are especially useful when the recombinant protein is uniformly labeled with

UCSF MISSION BAY CONTRA COSTA UNIVERSITY LIBRARY

^{15}N; for example, ^{1}H-^{15}N correlation spectra are excellent protein conformational fingerprints. (A conformational fingerprint is related to the overall tertiary structure as a classical two-dimensional peptide map is related to the primary sequence.) CD spectra in the far-UV region (240 to 180 nm) can be used to estimate secondary structure, and spectra in the near-UV region (350 to 240 nm) give information on the tertiary structure. When interpreting CD spectra, it should be borne in mind that a protein can have an apparently normal far-UV CD spectrum, yet have a near-UV spectrum characteristic of an unfolded protein. This is characteristic of partially folded protein or so-called molten globules (*UNIT 6.4*).

Apart from the analysis of recombinant proteins that contain the native sequence, mutant proteins are produced for various structure-function studies. It is almost mandatory for some conformational analyses to be made on these proteins so that functional differences can be related to specific chemical differences (i.e., different amino acid side chains) rather than merely being explained as due to conformational perturbation. For an example of a spectral-based analysis of site-directed mutants, see Mulkerrin and Cunningham (1993).

The folding question is also important for proteins prepared by other expression systems (e.g., yeast and baculovirus). Although these systems offer the advantage of providing proteins with the correct post-translational modifications and thus by supposition the correct structure, it is critical to demonstrate that the recombinant product has the same secondary and tertiary structures as the authentic protein. Incorrect disulfide bonding is sometimes encountered in these systems and will introduce conformational heterogeneity.

When analyzing protein conformation, it is a mistake to depend on any one physicochemical technique, especially the spectroscopic methods. The more corroborative data that is accumulated using different techniques and approaches, the more reliance can be placed on any conclusions.

For spectroscopic analysis, the ideal sample is protein freshly purified by all titration. The peak fractions of a symmetrically shaped elution peak are tested using the column elution buffer as reference material. If rapid buffer exchange is required, for example to remove excess salt prior to far-UV CD measurements, the sample (0.1 to 1.0 ml) can be applied to a prepacked column (1.5-cm diameter × 5 cm) of Sephadex G-25 (PD-10 columns from Pharmacia Biotech) and eluted in <5 min.

High-Resolution Structure

The ultimate test of a protein's structural integrity and conformational homogeneity is determination of its three-dimensional structure. For X-ray structure determination, crystals must be obtained that diffract to high resolution (<3 Å). The fact that the protein can be crystallized is in itself a fairly good indication that the preparation is both physically and chemically homogeneous. For reviews of the methodologies involved in protein crystallization and X-ray diffraction, see McRee (1993) and Ducruix and Giege (1992).

Structural determination of proteins in solution can be achieved by NMR spectroscopy. For small proteins (<10 kDa), conventional homonuclear two-dimensional methods have been used. For larger proteins, three- and four-dimensional heteronuclear methods are required to resolve the spectral complexity due to the larger numbers of protons. For this purpose, proteins expressed in *E. coli* are uniformly labeled with ^{15}N and/or ^{13}C as mentioned in *UNIT 5.3*. Structures of medium-sized proteins (~15 to 20 kDa) can be determined at high resolution using the current methodologies. For further details see references listed in Table 7.1.1 and Table 7.1.2.

Biological Activity and Immunogenicity

The biological activity of the isolated protein should be carefully studied. When possible, the activity should be compared to that of the authentic protein. For an apparently pure protein, a lower activity is often a clue that the protein has not folded correctly, is partially aggregated, or has been improperly post-translationally modified. A low specific activity may also indicate the presence of other proteinaceous contaminants that have been undetected by the analytical methods used. It is of course essential to accurately determine protein concentrations (Chapter 3) in order to make realistic comparisons between samples.

The main problem with all types of activity measurement is to know what proportion of the preparation is active. For an enzyme, active-site titration can give this information (Fersht, 1985). However, a small proportion (<5%) of any preparation may be inactive due to denaturation upon storage and standing, and this would not be detected by most of the spectroscopic methods referred to above. The presence

ECU IASHG CENTRAL UNIVERSITY LIBRARY

of small amounts of misfolded protein may normally be of no consequence, the main requirement being that the protein have similar biological activity to the standard. (The standard, especially if it is an authentic protein, will probably be even less well characterized than the recombinant protein).

If, on the other hand, the protein is to be used for in vivo studies, misfolded protein may be immunogenic and, thus, must be considered as a contaminant even though it may be chemically indistinguishable from the correctly folded protein. This is of special concern for therapeutic proteins, especially those repeatedly administered (for review, see Walsh and Headon, 1994).

The ability of antibodies to detect subtle conformational change has long been known (Goldberg, 1991). Conformational changes in local regions of the protein can be measured by determining changes in affinity. Furthermore, conformational differences between samples can be detected where spectroscopic and, more importantly, activity measurements, do not show a difference (Hattori et al., 1993).

BASIC PLAN OF ACTION

The task tree presented in Figure 7.1.1 is a suggested path for fully characterizing recombinant proteins. Some of the analytical techniques overlap in the type of information they provide. In order to avoid running multiple tests, we recommend using the high-technology methods, such as HPLC, mass spectrometry, circular dichroism, NMR, and fluorescence spectroscopy. However, if these are not available, good information can still be extracted from the more traditional approaches, keeping in mind that they are less resolving and therefore require complementary analytical strategies.

ANALYSIS OF N-TERMINAL HETEROGENEITY OF RECOMBINANT PROTEINS PRODUCED IN *E. COLI*

When a purified recombinant protein presents as a single band on an SDS-PAGE gel, gives a satisfactory amino acid analysis, gives the expected N-terminal amino acid sequence upon Edman degradation, shows the correct

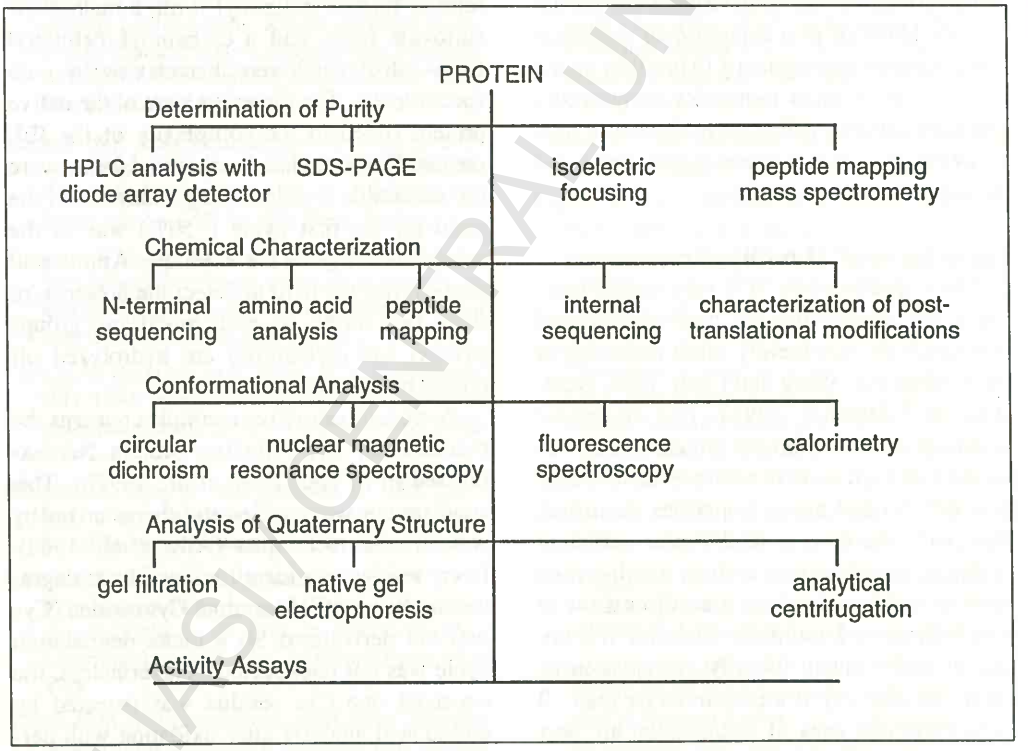


Figure 7.1.1 Task tree graphically representing the five steps (branches) required to fully characterize a recombinant protein. The first step is to determine the purity of the protein, which can be done by a variety of methods; the choice will depend on the availability of instrumentation and the degree of purity desired. This is followed by chemical characterization of the protein. Specific methods for the third and fourth steps, determination of protein conformation (folding) and structure, are listed in in Table 7.1.1 and Table 7.1.2. The last step, determining the activity of the protein, will require unique assays developed specifically for the protein of interest.

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

disulfide bond linkages, and is biologically active, most people are quite satisfied. However, insufficient resolution on SDS-PAGE, imprecise amino acid analysis (4% at best) and bioassays, and nonquantitative yield (particularly on the first cycle) during Edman degradation may all combine to mask protein heterogeneity.

High-resolution, charge-sensitive techniques, such as two-dimensional gel analysis involving isoelectric focusing (*UNIT 10.4*) and capillary electrophoresis, may reveal quite extensive heterogeneity in proteins that seem pure by the criteria described above. This heterogeneity may be due to deamidation of Asn or Gln residues—some of which, because of their position in the polypeptide chain, may be particularly susceptible to hydrolysis of their side chain amide—but other explanations are possible. For example, the carboxy terminus may be "ragged" (i.e., longer or shorter polypeptide chains may be present, with not all molecules terminating at the same residue). A ragged N-terminus is normally detected by Edman degradation, but other forms of post-translational modification at the N-terminus can go unnoticed because the amino group is blocked and so not accessible to the Edman chemistry. Various kinds of post-translational modification can create heterogeneity (when they do not affect all the protein molecules in question), and some can affect all of the molecules in such a way that the product is homogenous but is not the expected protein product.

Importance of Mass Spectrometry

Mass spectrometry is a very useful technique for identifying all post-translational modifications, not merely those occurring at the N-terminus (Wang and Chait, 1994; Bradshaw and Stewart, 1994). The molecular weight of the recombinant protein can be obtained with high accuracy and precision (better than 0.01%) and minor impurities identified. This performance is normally quite sufficient to detect modifications such as missing residues or additional groups (see discussion of Two Instructive Examples). However, it is not usually sufficient to identify deamidation (1 mass unit change) of a protein larger than 10 kDa. Particular sites of deamidation are best located by peptide mapping, and *Staphylococcus aureus* V8 endoprotease Glu-C and *Pseudomonas fragi* endoprotease Asp-N are useful in this regard because they can cleave at the sites of Glu and Asp residues, respectively (*UNIT 11.1*). If deamidation has occurred to a small extent at many sites, the problem of locating such sites becomes much more difficult.

Ideally, a protein should be purified to a single species prior to characterization, but this is not always possible. Once it is known that a portion of the sample has different charge (by IEF) or mass (by mass spectrometry), attempts should made to identify the site and nature of the modification by peptide mapping (*UNIT 11.6*). As is evident from the two examples that follow, and from *UNIT 11.6*, mass spectrometry has an essential role to play in the characterization of recombinant proteins.

Two Instructive Examples

A simple but instructive example is that provided by interleukin 5 (IL-5). Recombinant human IL-5 expressed in *E. coli* is a single band on SDS-PAGE but shows several bands on IEF (Rose et al., 1992a). Electrospray ionization mass spectrometry showed a major component with the expected molecular weight (observed, 13282.5 ± 2; calculated, 13280.4), but also showed evidence of the presence of species of slightly greater molecular weight. Peptide mapping led to isolation of four forms of the N-terminal peptide—the expected (methionyl) form, a formyl-methionyl form, a methionyl-sulfoxide form, and a carbamoyl-methionyl form—all of which were characterized by mass spectrometry. The dimeric nature of the native protein added to the complexity of the IEF results. The formyl and carbamoyl forms were not accessible to Edman degradation, and the yield on the first cycle (~50%) was in the acceptable range for the technique. Amino acid analysis did not help to detect the presence of these two forms, as both modifying groups (formyl and carbamoyl) are hydrolyzed off during hydrolysis.

A second instructive example concerns the recombinant DNA-binding protein Ner expressed in *E. coli* (Rose et al., 1992b). This small protein was purified and characterized by conventional techniques (Allet et al., 1988). Every residue was identified by Edman degradation except the N-terminal Cys residue (Cys was not derivatized, so a blank degradation cycle was not unexpected). Nevertheless, the expected one Cys residue was detected by amino acid analysis after oxidation with performic acid. It was only when mass spectrometry showed that the protein was too heavy by 70 mass units that a problem was recognized. After much effort, involving peptide mapping, mass spectrometry, microchemistry, and two-dimensional NMR spectroscopy, it turned out

CUA CITY OF HOPE NATIONAL MEDICAL CENTER LIBRARY

that the N-terminal residue was modified with pyruvate, but in a way not involving the carboxyl group of pyruvate (Rose et al., 1992b). This modification was unstable in mild acid and in mild base (although stable at physiological pH), which explained why it had been missed when the protein was analyzed by Edman degradation and amino acid analysis. Thus, although the protein was blocked (N-terminally modified), the modifying group became detached during the first coupling step of the Edman chemistry and so was invisible to the technique.

LITERATURE CITED

Ackers, G.K. 1970. Analytical gel chromatography of proteins. *Adv. Protein Chem.* 24:342-443.

Allet, B., Payton, M., Mattaliano, R.J., Gronenborn, A.M., Clore, G.M., and Wingfield. P.T. 1988. Purification and characterization of the DNA-binding protein ner of bacteriophage mu. *Gene* 65:259-268.

Ballery, N., Desmadril, M., Minard, P., and Yon, J.M. 1993. Characterization of an intermediate in the folding pathway of phosphoglycerate kinase: Chemical reactivity of genetically introduced cysteinyl residues during the folding process. *Biochemistry* 32:708-714.

Bradshaw, R.A. and Stewart, A.E. 1994. Analysis of proteins modifications: Recent advances in detection, characterization and mapping. Measurement as a tool for studying proteins. *Curr. Opin. Biotechnol.* 5:85-93.

Bristow, A.F. 1989. Purification of proteins for therapeutic use. *In* Protein Purification Applications: A Practical Approach (E.L.V. Harris and S. Angal, eds.) pp. 29-44. IRL Press, Oxford.

Bussian, B.M. and Sander, C. 1989. How to determine protein secondary structure in solution by Raman spectroscopy: Practical guide and test case DNase I. *Biochemistry* 28:4271-4277.

Carr, S.A., Hemling, M.E., Bean, M.F., and Roberts, G.D. 1991. Integration of mass spectrometry in analytical biotechnology. *Anal. Chem.* 63:2802-2824.

Chaiken, I., Rose, S., and Karlsson, R. 1992. Analysis of macromolecular interactions using immobilized ligands. *Anal. Biochem.* 201:197-210.

Chen, R.F., Endelhoch, H., and Steiner, R.F. 1969. Fluorescence of proteins. *In* Physical Principles and Techniques of Protein Chemistry (S.J. Leach, ed.) pp. 171-244. Academic Press, New York.

Clore, G.M. and Gronenborn, A.M. 1992. Methods of structural analysis of proteins. Part 2—Nuclear magnetic resonance. *In* Protein Engineering: A Practical Approach (A.R. Rees, M.J.E. Sternberg, and R. Wetzel, eds.) pp. 33-56. IRL Press, Oxford.

Clore, G.M. and Gronenborn, A.M. 1994. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* 239:349-363.

Cooper, A. and Johnson, C.M. 1994. Microcalorimetry; differential scanning calorimetry; and isothermal titration calorimetry. *Methods Mol. Biol.* 22:109-150.

Cunningham, B.C. and Wells, J.A. 1993. Comparison of a structural and functional epitope. *J. Mol. Biol.* 234:554-563.

Donovan, J.W. 1969. Ultraviolet absorption. *In* Physical Principals and Techniques of Protein Chemistry, Part A (S.J. Leach, ed.) pp. 101-170. Academic Press, New York.

Donovan, J.W. 1973. Ultraviolet difference spectroscopy—new techniques and applications. *Methods Enzymol.* 27:497-548.

Ducruix, A. and Giege, R. (eds.). 1992. Crystallization of Nucleic Acids and Proteins: A Practical Approach. IRL Press, Oxford.

Eftink, M.R. and Ghiron, C.A. 1981. Fluorescence quenching studies with proteins. *Anal. Biochem.* 114:199-227.

Englander, S.W. and Kallenbach, N.R. 1984. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16:521-655.

FDA (Food and Drug Administration). FDA Handbook. Points to Consider in the Production and Testing of New Drugs and Biologicals Produced by Recombinant DNA Technology. Office of Biologics Research and Review, Center for Drugs and Biologics, Food and Drug Administration, Bethesda, Md.

Fersht, A. 1985. Enzymes: Structure and Mechanism, 2nd Edition. Freeman, New York.

Figuet, B., Djavadi-Ohaniance, L., and Goldberg, M.E. 1989. Immunological analysis of protein conformation. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 287-310. IRL Press, Oxford.

Garnick, R.L., Rose, M.J., and Baffi, R.A. 1991. Characterization of proteins from recombinant DNA manufacture. *In* Drug Biotechnology Regulation: Scientific Basis and Practices. (Y.-Y. H. Chiu and J.L. Gueriguian, eds.) pp. 263-313. Marcel Decker, New York.

Geisow, M.J. 1991. Characterizing recombinant proteins. *Bio/Technology* 9:921-924.

Goldberg, M.E. 1991. Investigating protein conformation, dynamics and folding with monoclonal antibodies. *Trends Biochem. Sci.* 16:358-362.

Goldenberg, D.P. 1989. Analysis of protein conformation by gel electrophoresis. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 225-250. IRL Press, Oxford.

Hampton Research Macromolecular Crystallization Reagent Kits: Crystal Screen I and II. Hampton Research, Calif.

Harding, S.E. 1994. Sedimentation velocity; sedimentation equilibrium. *Methods Mol. Biol.* 22:61-83.

BOULDER CENTRAL UNIVERSITY LIBRARY

Harding, S.E. 1994. Classical light scattering and dynamic light scattering. *Methods Mol. Biol.* 22:61-83.

Haris, P.I. and Chapman, D. 1992. Does Fourier-transform infrared spectroscopy provide useful information on protein structures? *Trends Biochem. Sci.* 17:328-333.

Haris, P.I. and Chapman, D. 1994. Analysis of polypeptide and protein structures using Fourier-transform infrared spectroscopy. *Methods Mol. Biol.* 22:183-202.

Hattori, M., Ametani, A., Katakura, Y., Shimizu, M., and Kaminogawa, S. 1993. Unfolding/refolding studies on bovine β-lactoglobulin with monoclonal antibodies as probes. *J. Biol. Chem.* 268:22414-22419.

Haugland, R.P. 1994. Handbook of Fluorescent Probes and Research Chemicals, 5th Ed. Molecular Probes, Eugene, Oreg.

Johnson, W.C. 1990. Protein secondary structure and circular dichroism: A practical guide. *Proteins Struct. Funct. Genet.* 7:205-214.

LeGendre, N., Mansfield, M. Weiss, A., and Matsudaira, P. 1993. Purification of proteins and peptides by SDS-PAGE. *In* A Practical Guide to Protein and Peptide Purification for Microsequencing (P. Matsudaira, ed.) pp. 74-101. Academic Press, San Diego.

McRee, D.E. 1993. Practical Protein Crystallography. Academic Press, San Diego.

McRorie, D.K. and Voelker, P.J. 1993. Self-associating systems in the analytical ultracentrifuge. Vol. II, Analytical Ultracentrifugation Series. Beckman Instruments, Fullerton, Calif.

Mercier, J.-C., Grosclaude, F., and Ribadeau-Dumas, B. 1971. Primary structure of bovine S1 casein. Complete sequence. *Eur. J. Biochem.* 23:41-51.

Mulkerrin, M.G. and Cunningham, B.C. 1993. Spectral analysis of site-directed mutants of human growth hormone. *In* Protein Folding: In Vivo and In Vitro (J.L. Cleland, ed.) pp. 240-253. ACS Symposium Series 526. American Chemical Society, Washington, DC.

Pappin, D.J.C., Hojrup, P., and Bleasby, A.J. 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* 3:327-333.

Pennica, D., Kohr, W.J., Fendly, B.M., Shire, S.J., Raab, H.E., Borchardt, P.E., Lewis, M., and Goeddel, D.V. 1992. Characterization of a recombinant extracellular domain of the type I tumor necrosis factor receptor. *Biochemistry* 31:1134-1141.

Phillies, G.D. 1990. Quasielastic light scattering. *Anal. Chem.* 62:1049-1057.

Price, N.A. and Johnson, C.M. 1989. Proteinases as probes of conformation of soluble proteins. *In* Proteolytic Enzymes: A Practical Approach (R.J. Beynon and J.S. Bond, eds.) pp. 163-179. IRL Press, Oxford.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E., and Razgulyaev, O.I. 1990. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* 262:20-24.

Ralston, G. 1993. Introduction to analytical ultracentrifugation. Vol. 1, Analytical Ultracentrifugation Series. Beckman Instruments, Fullerton, Calif.

Rose, K., Anderegg, R., Wells, T.N., and Proudfoot, A.E. 1992a. Human interleukin-5 expressed in *E. coli* has N-terminal modifications. *Biochem. J.* 286:825-828.

Rose, K., Simona, M. G., Savoy, L.-A., Green, B. M., Gronenborn, A.M., Clore, G.M., and Wingfield, P.T. 1992b. Pyruvic acid is attached to the amino terminus of the recombinant-DNA derived DNA-binding protein ner of bacteriophage mu. *J. Biol Chem.* 267:19101-19106.

Scoble, H.A. and Martin, S.A. 1990. Characterization of recombinant proteins. *Methods Enzymol.* 193:519-536.

Shire, S.J. 1992. Determination of the molecular weight of glycoproteins by analytical ultracentrifugation. Beckman Technical Information, Publication DS-837. Spinco Business Unit, Palo Alto, Calif.

Strickland, E.H. 1974. Aromatic contributions to circular dichroism spectra of proteins. *Crit. Rev. Biochem.* 2:113-175.

Sturtevant, J.M. 1987. Biochemical applications of differential scanning calorimetry. *Ann. Rev. Phys. Chem.* 38:463-488.

Surewicz, W.K., Mantsch, H.H., and Chapman, D. 1993. Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment. *Biochemistry* 32:389-393.

Thomas, D., Schultz, P., Steven, A.C., and Wall, J.S. 1994. Mass analysis of biological macromolecular complexes by STEM. *Biol. Cell* 80:181-192.

Van Holde, K.E. 1975. Sedimentation analysis of proteins. *In* The Proteins, 3rd Edition, Vol. 1 (H. Neurath, ed.) pp. 225-291. Academic Press, N.Y.

Varley, P.G. 1994. Fluorescence spectroscopy. *Methods Mol. Biol.* 22:203-218.

Walsh, G. and Headon, D.R. 1994. Therapeutic proteins: Special considerations. *In* Protein Biotechnology. pp. 119-162. John Wiley & Sons, Chichester, U.K.

Wang, R. and Chait, B.T. 1994. High-accuracy mass measurement as a tool for studying proteins. *Curr. Opin. Biotechnol.* 5:77-84.

Wang, R., Chait, B.T., and Kent, S.B.H. 1994. Protein ladder sequencing towards automation in techniques. *In* Protein Chemistry V (J.W. Crabb, ed.) pp. 19-26. Academic Press, New York.

Wetlaufer, D.B. 1962. Ultraviolet spectra of proteins and amino acids. *Adv. Protein Chem.* 17:303-390.

FGU IASI LIBRARY

Wettenhall, R.E. and Cohen, P. 1982. Isolation and characterization of cyclic AMP–dependent phosphorylation sites from rat liver ribosomal protein S6. *FEBS Lett.* 140:263-269.

Williams, R.W. 1986. Protein secondary structure analysis using Raman amide I and III spectra. *Methods Enzymol.* 130:311-331.

Wood, S.P. 1989. Purification for crystallography. *In* Protein Purification Applications. A Practical Approach (E.L.V. Harris and S. Angal, eds.) pp. 45-58. IRL Press, Oxford.

Yang, J.T., Wu, C.-S., and Martinez, H. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol.* 130:208-297.

KEY REFERENCES

Cantor, C.R. and Schimmel, P.R. 1980. Biophysical Chemistry. Part II: Techniques for the Study of Biological Structure and Function. Freeman, San Francisco.

Fairly advanced text that includes descriptions of the spectroscopic and ultracentrifugation methods.

Freifelder, D. 1982. Physical Biochemistry: Applications to Biochemistry and Molecular Biology. Freeman, N.Y.

Excellent introductory text with especially strong sections on spectroscopy and centrifugation. Read this book first.

Kyte, J. 1995. Structure in Protein Chemistry. Garland Publishing, N.Y.

Structural biology emphasis, with a clear and concise explanation of X-ray diffraction analysis. Excellent illustrations.

Campbell, I.D. and Dwek, R.A. 1984. Biological Spectroscopy. Benjamin/Cummings Company, Menlo Park, Calif.

Very good explanations of the theoretical principles, along with many practical applications.

Contributed by Nancy D. Denslow
University of Florida
Gainesville, Florida

Paul T. Wingfield
National Institutes of Health
Bethesda, Maryland

Keith Rose
University Medical Centre (CMU)
Geneva, Switzerland

BCU IASI // CENTRAL UNIVERSITY LIBRARY

Determining the Identity and Purity of Recombinant Proteins by UV Absorption Spectroscopy

Ultraviolet absorption spectrophotometry can be used to confirm the identity and, to a lesser extent, assess the purity of recombinant proteins and their peptide fragments. The near-UV (250 to 350 nm) absorbance spectrum of a protein is almost entirely a function of its aromatic amino acids—tryptophan, tyrosine, and phenylalanine (Fig. 7.2.1). Because each protein (gene product) has a unique amino acid sequence, the particular aromatic amino acid content of each protein results in a unique spectrum in the near-UV region. The highly specific microenvironment experienced by each aromatic residue in the three-dimensional protein matrix results in fine shifts in a protein's spectrum, which can be detected (Basic Protocol 1) and analyzed (Support Protocol). Thus, quality UV spectra serve as indicators of the structural integrity of proteins. In contrast to electrophoretic analyses (Chapter 10), UV spectroscopy allows the detection of nucleic acids, certain prosthetic groups, and metal ions which are frequently associated with proteins.

The unique UV spectral properties of proteins can in turn be used to assess their purity. This application is inherent in the use of a diode array detector to monitor the effluent

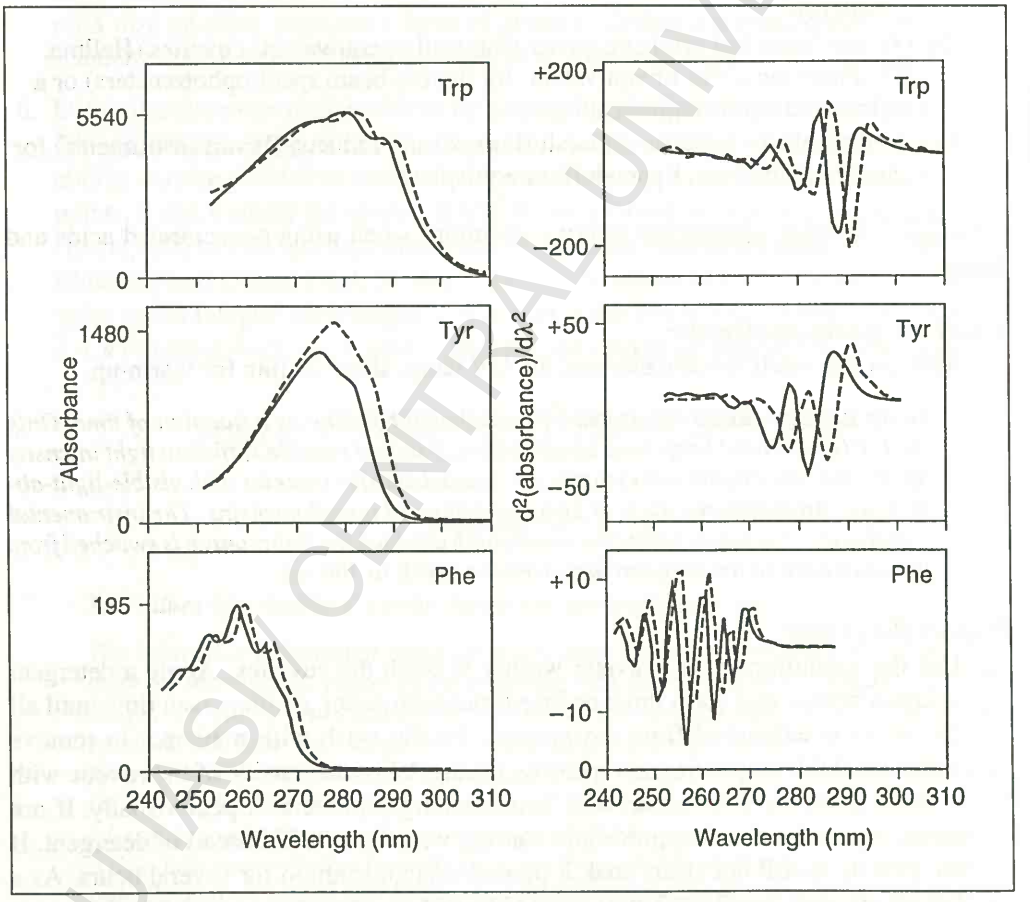


Figure 7.2.1 Normal (zero-order) and second-derivative spectra of aromatic model compounds: *N*-acetyl-L-tryptophanamide, *N*-acetyl-L-tyrosineamide, and *N*-acetylphenyl ethyl ester (solid lines) and "average" aromatic amino acid residues as determined using a set of globular, water-soluble proteins (dotted lines). Absorbance and second-derivative values correspond to 1 M solutions, 1-cm path length. The band positions in denatured proteins usually have intermediate positions.

Copyright © 1995 by John Wiley & Sons, Inc.

UNSW CENTRAL UNIVERSITY LIBRARY

from a high-performance liquid chromatography (HPLC) column. A protocol using this technique to assess the purity of recombinant proteins is presented in Basic Protocol 2.

BASIC PROTOCOL 1

ANALYSIS OF PROTEINS USING NEAR-UV SPECTROPHOTOMETRY

This procedure identifies proteins using near-UV spectrophotometry. Protein solution is placed in a UV-transparent quartz cuvette, absorption of light is measured as a function of wavelength, and the resulting spectrum is transferred either to a displaying device (e.g., CRT or printer) for inspection or to a computer that can employ a spreadsheet or other program for derivative calculations and/or quantitative analyses (see Support Protocol).

Materials

Detergent solution
Methanol
0.1 N hydrochloric acid
Acid/ethanol cleaning solution (see recipe)
10 N NaOH
4% holmium oxide (aqueous solution in 10% perchloric acid; see recipe)
Buffer identical to that containing the protein (reference standard)
Sample for analysis in solution at an appropriate concentration (e.g., 0.05 to 1 mg/ml; see below)

Vacuum-powered cuvette washer (NSG Precision Cells, Kontes Glass, or equivalent)
Two factory-matched synthetic quartz (Suprasil or equivalent) cuvettes (Hellma, NSG Precision Cells, or equivalent; for double-beam spectrophotometers) or a single cuvette (for single-beam use)
Gel-loading plastic pipet tips (Marsh Biomedical Products; Rainin Instruments) for a pipettor (Pipetman, Eppendorf, or equivalent)

CAUTION: Employ appropriate safety precautions when using concentrated acids and bases.

Ready the spectrophotometer

1. Turn on the spectrophotometer and the UV lamp; allow 20 min for warm-up.

   Lamp stability can be monitored by recording a baseline as a function of time. Only the UV (deuterium) lamp need be turned on; it should provide sufficient light intensity up to ~400 nm. Higher wavelengths are needed only for proteins with visible-light-absorbing chromophores, such as heme proteins and metalloproteins. The instrumental wavelength cross-over point (the wavelength at which the light source is switched from the deuterium to the tungsten lamp) must be set ≥ to 360 nm.

Prepare the cuvette

2. Use the vacuum-powered cuvette washer to wash the cuvettes. Apply a detergent solution briefly and wash three or four times with water, waiting each time until all the water is exhausted from the cuvette. Finally wash with methanol to remove water-insoluble impurities and speed up drying. Wipe the outside of the cuvette with a tissue (Kimwipe) or, better, with lens cleaning paper and inspect visually. If not clean, repeat the cleaning procedure starting with 0.1 N HCl instead of detergent. If the cuvette is still not clean, soak it in acid/ethanol solution for several hours. As a last resort, soak briefly (20 min) in 10 N NaOH.

   Cuvettes should always be cleaned immediately after use. Denatured proteins may often be removed by soaking overnight in a mild pepsin solution. Prolonged contact with the NaOH may partially dissolve quartz faces of the cuvette.

CSU IASI CENTRAL UNIVERSITY LIBRARY

Set up appropriate instrument parameters

3. Identify the cross-sectional area of the light beam (from the instrument's instruction manual) and determine the correct height to which the cuvettes should be fitted.

   The light beam may be as large as 1 × 2 cm. The sample volume must be large enough so that the meniscus is completely above the beam. The light must not impinge on the air-solution interface. If the beam is wider than the portion of the cuvette containing the solution, the remaining portions of the cuvette walls should be masked by a light-absorbing coat; special blackened cuvettes are available for this purpose. In some instruments the height of the beam can also be reduced; do so if sample availability is a problem. The cuvettes must be positioned vertically; any deviations may alter the effective path length and reduce the precision of measurements.

4. Adjust the wavelength (scanning) range to 240- to 360-nm. For diode-array spectrophotometers, use 10-sec measurement times. For conventional scanning spectrophotometers use bandwidths of 1 or 2 nm, time constants (response times) of 2 sec (medium) or less, and scanning speeds of 60 nm/min or less. Set the data point interval to 1 nm (scanning instruments only).

   Use the same settings if spectra are to be compared to previously collected ones.

Measure the spectrum of a reference standard

5. Measure the baseline (background) spectrum with no cuvettes in the sample holders. Insert a sealed cuvette containing 4% holmium oxide (aqueous solution in perchloric acid; Weidner et al., 1985) into the sample holder. Measure the spectrum and save it on a disc or other permanent form of storage. Collect a second spectrum using a scanning speed approximately one-half of that initially used.

6. Determine the exact peak positions by calculating the first-derivative spectrum in the 280- to 290-nm range. If the spectrophotometer has a built-in derivative calculation ability, choose derivative order 1, polynomial degree 2, and a window of five data points. If not, transfer the spectra in ASCII (text) format (or type it) into a spreadsheet (Excel, LOTUS-123, or equivalent) and perform the calculation using Equation 7.2.1 (Savitzky and Golay, 1964; Steiner et al., 1972), where FD(λ) is the first-derivative value at the integral wavelength λ, $A(\lambda \pm n)$ is the absorbance values at wavelength $\lambda + n$ (where $n = -2, -1, 1$, and 2), and k is the data point interval (in nanometers).

$$FD(\lambda) = \frac{-2[A(\lambda-2)] - A(\lambda-1) + A(\lambda+1) + 2[A(\lambda+2)]}{10 \times k}$$

Equation 7.2.1

   The values of n double if a 2-nm data point spacing is employed.

   The derivative calculation using an Excel spreadsheet is performed as follows: (1) copy the absorbance values of the spectrum into the first column (typically A1:A100 range); (2) fill second column (B1:B100) with the corresponding wavelength values, entering the first wavelength and using Edit/Fill/Series commands; (3) enter the formula in the cell of the third column, which corresponds to A(λ) absorbance value (B3)—e.g., for Equation 7.2.1 and k=1, cell C3 contains =(−2×A1−A2+A4+2×A5)/10 (derivative values cannot be calculated for first two and last two data points because a five-data point window is used); (4) copy the content of that cell (the formula) down onto the entire column (B3:B98) using Edit/Copy and Edit/Paste commands; and (5) use the Chart Wizard option to plot the resulting spectrum.

7. Note where the spectra cross the zero line (change sign from positive to negative) and calculate the exact positions of these intercepts using Equation 7.2.2, where λ is

BOULDER / CENTRAL UNIVERSITY LIBRARY

the wavelength prior to the intersection point, k is the data point interval (in nanometers), and |FD(λ)| and |FD(λ + 1)| are the absolute values of the first derivative at the integral wavelengths λ (before the intersection point) and λ + 1 (after the intersection point).

$$\text{peak position (nm)} = \lambda + k\frac{|FD(\lambda)|}{|FD(\lambda)| + |FD(\lambda+1)|}$$

Equation 7.2.2

This equation arises from simple geometric interpolation, illustrated for the second-derivative spectra in the inset of Figure 7.2.2B.

Note that the resulting intersection point is the position of the peak near 287 nm and serves as a reference point for further determination of exact peak positions of the protein spectra.

8. Calculate the intersection point for the second spectrum collected using a slower scanning speed (scanning instruments only); if the position is different, decrease the time constant (response time) or use a slower scanning speed until the positions of the intersections become independent of these parameters.

Test for the proper response time only with the first spectra acquired. Once established, the settings can then be used with confidence. The 287-nm holmium oxide peak is found at 287.18 nm (for 1-nm bandwidths) and at 287.47 nm (for 2- and 3-nm bandwidth) on an instrument extensively verified by measurements of mercury and deuterium emission lines (Weidner et al., 1985). Commercial instruments will deviate somewhat from this value. Correct the wavelength results for proteins for the difference observed. Diode-array instruments (see Basic Protocol 2) will typically require such calibration once a year; scanning instruments should be calibrated frequently depending on use. If a holmium oxide standard is not readily available, Trp (monomeric amino acid) or a related compound can be used for reference purposes.

Measure the protein spectrum

9. Identify the number of light beams by looking at the number of sample holders and their positions in the sample compartment. In a double-beam spectrophotometer, use two matched cuvettes and perform an initial background (baseline) measurement with just buffer (identical to that in which the protein sample is dissolved) in both cuvettes. Then substitute protein solution for the buffer in the sample position cuvette, leaving the reference buffer in the other cuvette. Acquire the sample spectrum and subtract the previously measured background spectrum if necessary. In a single beam spectrophotometer, use a single cuvette. Measure the background with buffer in the cuvette. Then remove the buffer, measure the spectrum of the unfiltered protein solution using the same cuvette in the same orientation, and subtract the background spectrum (baseline).

Sample can be removed from the cuvette using a pipettor with gel-loading plastic pipet tips. Do not scratch the inner optical surface of the cuvette. Many proteins, especially when partially unfolded, strongly adhere to quartz and are not removed by brief detergent or organic solvent washes. If UV absorption peaks at 270- to 280-nm are detected in the baseline, clean the cuvette again (step 2).

10. Note the absorbance at the protein peak centered between 275 and 282 nm. If the absorbance value is >1.0 at the maximum, check the range of linearity of the instrument in use (from the instruction manual). If the absorbance value observed is above this range, dilute the sample or decrease the path length. If a fluorescence-type 2-5 × 10–mm cell is used, turn the cuvette sideways to reduce the path length. A new baseline should be obtained. If the absorbance falls below 0.2, carefully observe the

VJU IASI CENTRAL UNIVERSITY LIBRARY

noise levels and use longer acquisition times (both for the baseline and the actual spectrum).

11. Save the spectrum in a proper format for the particular instrument employed or transfer the spectrum in an ASCII (text) format to a disc for retrieval by a spreadsheet for further analysis. If this is not feasible, print or manually record absorbance values as a function of wavelength.

Most protein solutions containing 1 mg/ml protein have an absorbance near 1.1 ± 0.5 in the 280-nm region. If an estimate of the concentration of the solution is available, the sample can be diluted or concentrated to obtain an absorbance in the 0.2 to 1.0 range. Alternately, make a preliminary measurement at a single wavelength (e.g., 280 nm) to obtain this information. Note that with modern spectrophotometers very low values of absorbance (<0.01) can be accurately measured; however, care must be taken to avoid adsorption artifacts, as a substantial portion of the protein may now reside on the interior surface of the cuvette. This can be checked by obtaining a spectrum of the empty cuvette after carefully removing the protein solution.

INTERPRETING PROTEIN NEAR-UV SPECTRA

SUPPORT PROTOCOL

Normal (zero-order) spectra are spectra of absorbance versus wavelength as measured by a spectrophotometer without derivative calculation. Different regions of the spectrum contain information about various aspects of the structure of the protein. The 320- to 350-nm region may contain information about aggregation state (due to light scattering) or the presence of UV-absorbing prosthetic groups; the 250- to 300-nm region provides information about the aromatic amino acids and their environment (Fig. 7.2.1 and Table 7.2.1). Numerical calculation of the second derivative of the spectrum minimizes the effect of broad spectral components (such as light scattering or anion absorption), exaggerating fine structure originating from absorption by aromatic rings. This method is particularly useful for precise determination of tertiary structure–dependent band positions as well as for quantitative analysis.

Correct for light scattering

1. Examine the spectral region between 320 and 350 nm (Fig. 7.2.2 contains an example for a typical protein).

Significant "apparent absorbance" of gradually, almost linearly decreasing intensity with increasing wavelength indicates that a portion of the light beam did not reach the detector because it was scattered. Absorbance that falls to zero at 350 nm or does not gradually decline (Fig. 7.2.2) is probably true absorbance arising from prosthetic groups or contaminants.

2. Use Equation 7.2.3 to calculate optical density (OD) due to light scattering, where a and b are constants.

$$\log OD = a \log \lambda + b$$

Equation 7.2.3

If the instrument's software cannot make such a correction, use a spreadsheet (Microsoft Excel, LOTUS-123, or equivalent; see Basic Protocol 1, step 6 annotation, for an example) to generate the light-scattering curve using optical density values at 320 nm and 350 nm. Calculate the light-scattering correction from Equation 7.2.4, where $m = 64.32 - 25.67 \log \lambda$.

$$A_c = 10^{[(m+1) \times \log OD_{320} - m \times \log OD_{350}]}$$

Equation 7.2.4

BCU RSI / CENTRAL UNIVERSITY LIBRARY

If this is not feasible, calculate the light-scattering OD at 260 nm (m = 2.31) and 280 nm (m = 1.50) with a calculator. If significant light scattering is detected, remeasure the sample after filtration through a 0.1 to 0.45-micron diameter cut-off filter. The difference of protein concentration before and after filtration will indicate the fraction of aggregated protein. Compare the second-derivative spectra to determine whether the initial sample was homogenous.

Note that optical density (OD) refers to the total extinction of light produced by all processes (including light scattering) whereas absorbance (A) refers only to the actual absorption process. The coefficient a varies from –4 (the ideal case of scattering from a small particle, not larger than one-tenth the wavelength of the light; this situation is known as Rayleigh scattering) to +2.2 for submicron-sized aggregates (Mie theory, in which the scattering intensity becomes a complex function of molecular size; Timasheff, 1966). Because the light-scattering contribution to the optical density is almost linear in the 300- to 350-nm region, a simple linear fit using a spreadsheet or even a crude graphic estimate can serve as a rudimentary measure of the correction necessary to obtain the correct absorbance in the 250- to 300-nm range.

Visually recognize the presence of the individual types of aromatic residues

3. Inspect the spectrum in the 250- to 300-nm range.

A small inflection point on the declining shoulder of the main peak around 290 nm (Fig. 7.2.2) originates from Trp residues; the sharper the inflection point, the more homogeneous and apolar the average microenvironment of these side chains. The distinctive spectral features of Tyr residues are usually masked and less pronounced as a consequence of the less-sharp fine structure of these residues, their broadening by the medium, and their greater heterogeneity from their usually greater frequency in proteins. The presence of the Tyr residues can be deduced, however, from the deviation of the overall spectral shape from that expected for pure Trp. The presence of Phe residues is usually manifested as a series of weak but distinctive bumps superimposed on the rising 250- to 270-nm region, which is dominated by more strongly absorbing Trp and Tyr side chains.

Table 7.2.1 Spectroscopic Properties of UV-Absorbing Amino Acids, Natural Ligands (Prosthetic Groups), and Nucleic Acids

Chromophore	λ (nm)[a]	ε (M^{-1} cm^{-1})[b]
Tryptophan (in native proteins)	280	5540
Tyrosine (in native proteins)	280	1480
Disulfide bond (glutathione)	280	134
N-Acetyl-L-tryptophanamide	280	5390
N-Acetyl-L-tyrosinamide	275	1390
N-Acetyl-L-tyrosinamide	280	1185
N-Acetyl-L-phenylalanine ethyl ester	257	195
Cu^{2+} (azurin)	781	320
FAD (pyruvic dehydrogenase)	460	1270
Fe^{3+}-heme (cytochrome *c*, reduced)	550	2770
FMN (amino acid oxidase)	455	1270
Retinal-Lys (rhodopsin)	498	4200
DNA (native, per base)	258	6600
RNA (per base)	258	7400

[a]Wavelengths listed indicate the peak maxima, except for tyrosine, *N*-Acetyl-L-tyrosinamide, and L-cystine at 280 nm.

[b]Molar absorptivity values of tryptophan and tyrosine in native proteins as well as oxidized glutathione (disulfide bond) at 280 nm are used for the calculation of molar extinction coefficients of proteins.

EDU IASI / CENTRAL UNIVERSITY LIBRARY

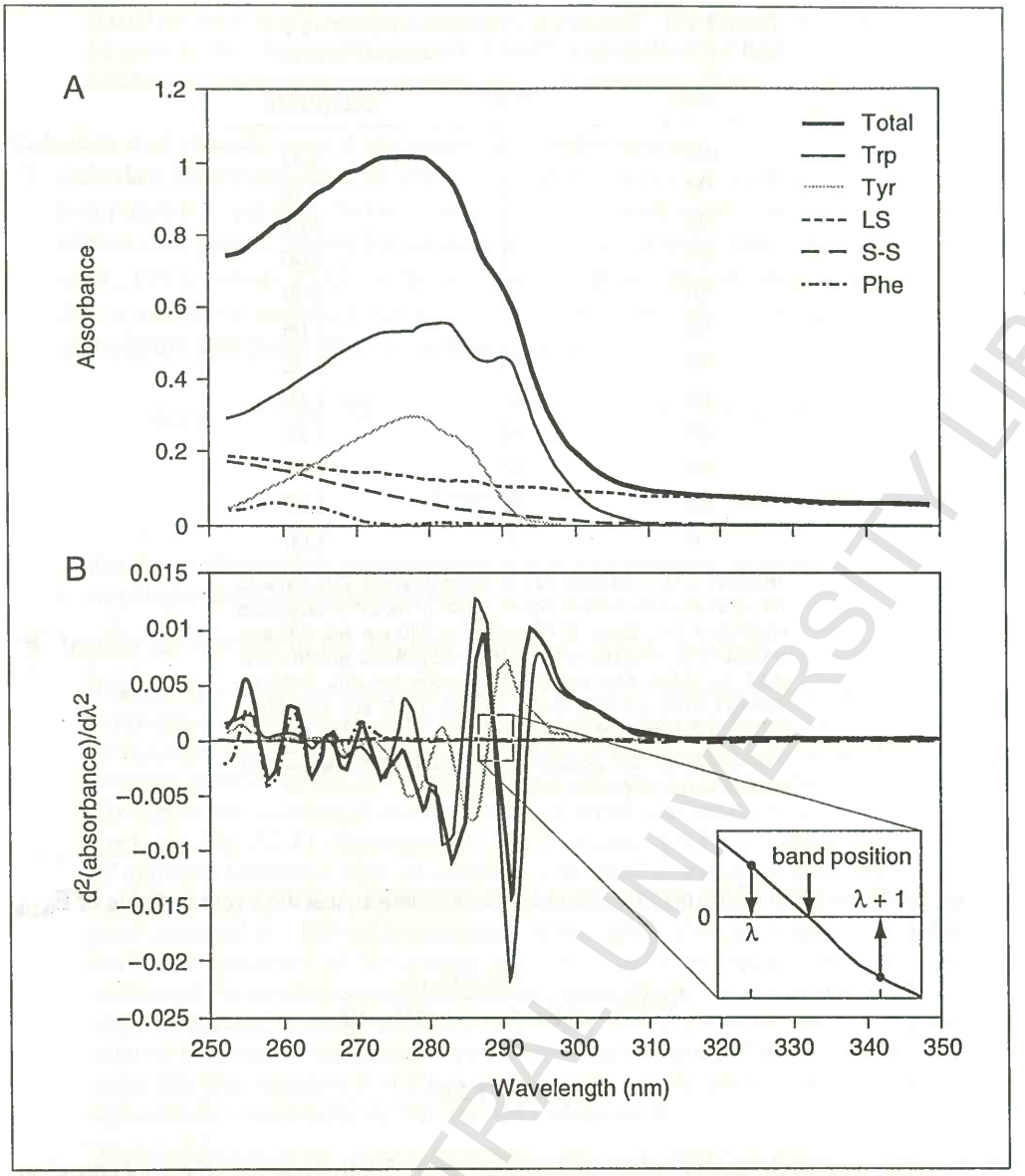


Figure 7.2.2 (**A**) Normal (zero-order) spectrum of a 100 µM solution of a hypothetical protein containing 1 Trp, 2 Tyr, 3 Phe, and 4 Cys (S-S) residues. The constituent spectral components and a typical light scattering (LS) contribution are also shown. (**B**) Calculated second-derivative spectra of **A**. The second-derivative spectra of cystine and light-scattering components are essentially zero. Inset: Enlargement of the marked segment illustrating the interpolation between adjacent second-derivative points that yields the precise position of the intersect with the wavelength axis.

Estimate the nucleic acid content

4. Calculate the ratio R of absorbances at 260 nm and 280 nm, $A_{260,c}$ to $A_{280,c}$ after correction for light scattering (Equation 7.2.4). Use Table 7.2.2 to estimate % weight of DNA (%N) in the sample (Glasel, 1995).

Estimate protein concentration in the sample

5. If a significant amount of DNA was detected (R>0.70), calculate the protein absorbance $A_{280,c,p}$ using % weight of protein and DNA found in Table 7.2.2, where $E_{0.1\%}$ is the extinction coefficient of the protein (per 1 mg/ml). If the amino acid composition is known, calculate $E_{0.1\%}$ using Equation 7.2.7 and divide the result (ε) by the

BOSTON UNIVERSITY / CENTRAL UNIVERSITY LIBRARY

Table 7.2.2 Theoretical R(260/280) Values vs %P and %N in Mixtures of Nucleic Acids and Proteins[a]

%P[b]	%N[b]	R(260/280)
100	0	0.57
99	1	0.70
98	2	0.81
97	3	0.90
96	4	0.99
95	5	1.06
90	10	1.32
80	20	1.59
60	40	1.81
40	60	1.91
20	80	1.97
0	100	2.00

[a]Amount of DNA as small as 1% can be detected. This is due to the approximately tenfold higher values of the DNA extinction coefficient (i.e., $E_{0.1\%}$ of 10 and 20 at 280 nm and 260 nm, respectively). Average value of R for 12 globular proteins was 0.57 +/– 0.06. The empirical equation for this table is: %N=(11.16×R–6.32)/(2.16–R) (Glasel, 1995; see also Manchester, 1995, for further comments).

[b]Abbreviations: %N, percent weight of DNA in the sample; %P, percent weight of protein in the sample.

molecular weight of the protein. For a crude estimation, use the average value of $E_{0.1\%} = 1.1$.

$$A_{280,c,p} = \frac{\%P \times E_{0.1\%}}{\%P \times E_{0.1\%} + \%N \times 10}$$

Equation 7.2.5

6. Use the spectral intensity values at 280, 320, and 350 nm to assess the concentration of protein based on the known amino acid composition of the protein using Equation 7.2.6 and Equation 7.2.7 (Mach et al., 1992b), where C is the protein concentration (molar), ε is the molar extinction coefficient (M^{-1} cm^{-1}) that would be obtained from a solution of 1 M concentration and a 1-cm path length (L), and N_{Trp}, N_{Tyr}, and $N_{S\text{-}S}$ are the numbers of Trp, Tyr, and disulfide bonds (cystine residues) in the protein. $A_{280,c}$ is the optical density at 280 nm (OD_{280}) corrected for light scattering (see Equation 7.2.4). When sufficient amount of protein is available, the extinction coefficient ε is usually determined by quantitative amino acid analysis (*UNIT 3.2*).

$$C = \frac{A_{280,c}}{\varepsilon L}$$

Equation 7.2.6

$$\varepsilon = 5540N_{Trp} + 1480N_{Tyr} + 134N_{S\text{-}S}$$

Equation 7.2.7

CU IASH CENTRAL UNIVERSITY LIBRARY

Based on studies of a series of proteins (Equation 7.2.7), the above relationships can be used to determine the extinction coefficients of proteins and, consequently, their aqueous concentration to an accuracy generally within 2%.

Calculate and visually inspect the second-derivative spectra

7. Calculate the second-derivative spectrum. If the instrument's software has derivative computation capability, choose derivative order 2, polynomial order 2, and a window of five data points. Otherwise use Equation 7.2.8 (Savitzky and Golay, 1964; Steiner et al., 1972), where SD(λ) is the second-derivative value at integral wavelength λ, $A(\lambda \pm n)$ is the absorbance value at $\lambda + n$ wavelength (where n = –2,–1, 0, 1 and 2), and k is the data point interval (in nanometers).

$$SD(\lambda) = \frac{2[A(\lambda-2)] - A(\lambda-1) - 2[A(\lambda) - A(\lambda+1) + 2[A(\lambda+2)]}{7 \times k}$$

Equation 7.2.8

See Basic Protocol 1, step 6 annotation, for an example. Note that the values of n double for 2-nm spectral bandwidth.

8. Inspect the features of the resultant second-derivative spectrum.

The negative peak centered near 290 nm (Fig. 7.2.2) originates from the Trp inflection point observed in the zero-order spectrum. Although its position is partially influenced by the overlapping Tyr spectrum, the position of the negative 290-nm peak reflects the average polarity of Trp microenvironments: the more hydrophobic the environment, the higher the wavelength (i.e., an apolar environment produces a "red shift" in this peak; see Fig. 7.2.1). The magnitude of this negative peak is diminished by a Tyr band of opposite (positive) sign. In contrast, Trp and Tyr bands of the same sign (both negative) amplify the composite second-derivative peak near 280 nm. The positive peak centered at ~285 nm is composed mostly of the Trp component, as the second-derivative spectrum of Tyr crosses the zero axis in this region. In some proteins individual Tyr residues have microenvironments different enough to produce 2- to 3-nm shifts and band overlap sufficient to obviate individual peaks in this region. The alternating positive and negative peaks between 250 and 270 nm are the result of the wave-like fine structure from Phe side chains seen in the zero-order spectrum and are affected only marginally by Trp and Tyr counterparts.

The magnitudes of the second-derivative peaks are proportional to the concentration of aromatic amino acids, revealing both the relative content of each type of aromatic residue and the protein concentration. The relative heights of the second-derivative peaks constitute a characteristic "fingerprint" by which purity and structural integrity of a protein can be recognized. Some instruments (e.g., Hewlett-Packard 8452) have software capable of comparing two spectra using linear regression and providing a quantitative "match factor." Unlike that for holmium oxide, the first derivative cannot be used for protein solutions because of residual nonzero values originating from light scattering, cystine, and other amino acid absorbance below 270 nm.

9. Calculate the positions of the intersections of the second derivative spectrum with the wavelength axis in the 285- to 295-nm region. Use Equation 7.2.9, where λ is the wavelength prior to the intersection point, k is the data point interval (in nanometers), and |SD(λ)| and |SD(λ + 1)| are the absolute values of the second derivative at the integral wavelengths λ (before the intersection point) and λ + 1 (after the intersection point).

$$\text{peak position (nm)} = \lambda + k\frac{|SD(\lambda)|}{|SD(\lambda)| + |SD(\lambda+1)|}$$

Equation 7.2.9

B. C. U. „M. Eminescu" IAŞI
CHIMIE

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

These values can be used as probes of protein structural integrity. In future manipulations or purification of the same protein, identical values of these parameters will indicate that the structure of the protein did not significantly change. The values of intersection points are typically reproducible within a 0.05-nm error margin using diode-array spectrophotometers. Scanning spectrophotometers require frequent wavelength calibration using standardized spectra to obtain this precision. The intersections of the second-derivative spectra of the holmium oxide standard determined using Equation 7.2.9 provide suitable reference points with which to compare results obtained at different times or using different instruments.

BASIC PROTOCOL 2

ANALYSIS OF PROTEINS USING AN HPLC WITH A DIODE-ARRAY DETECTOR

Purifying recombinant proteins relies on large-scale chromatographic procedures that generally lack the resolution of the analytical methods on which they are based. A common problem in the large-scale process is that a second protein (or some other compound, e.g., a detergent) may coelute with the protein of interest, contaminating the sample. High-performance liquid chromatography (HPLC) with a diode-array detector is a good method to determine the purity of a sample. By exquisitely regulating the pumps, this chromatography system minimizes flow problems that can otherwise interfere with peak detection, yielding highly reproducible gradients. Thus, HPLC provides the best chromatographic resolution of a protein.

A diode-array detector can be coupled to virtually any type of chromatographic separation, including reversed-phase, ion exchange (*UNIT 8.2*), and gel filtration (*UNIT 8.3*). The method assesses the purity of proteins or peptides equally well. In this protocol, the purity of a protein is determined using a C_4 reversed-phase column. The same protocol can be adapted to peptides by substituting a C_{18} reversed-phase column.

Materials

Buffer A: 0.1% (v/v) trifluoroacetic acid (TFA; reagent grade) in water (Milli-Q or equivalent purity)

Buffer B: 0.09% (v/v) TFA in 80% acetonitrile (HPLC grade)

C_4 reversed phase column and guard: (e.g., Vydac C_4, 250 × 4.6-mm-i.d. column, 5-μm particle size, 300-Å pore size for samples in the 0.5 to 5 nmol range or a 250 × 2.1-mm-i.d. column for samples in the 100 to 500 pmol range)

HPLC with a diode-array detector (e.g., Hewlett-Packard, Waters, Beckman)

1. Set scan limits on the diode array detector to cover the 200- to 300-nm region.

   Peptide bonds have strong absorption in the 200- to 230-nm region (wavelength maximum = 188 nm) and aromatic amino acids have specific maxima in the 240- to 300-nm region.

2. Set the primary sample wavelength at 215 nm, with secondary tracings being made at 254 nm and 280 nm. Set bandwidth at 4 nm.

   The sample wavelength is the wavelength at which absorbance is measured as a function of time for the whole run. The bandwidth can be set from 4- to 20-nm for a sample. A bandwidth of 4 nm for the primary wavelength of 215 nm means that the absorbance measurement will extend between 213 and 217 nm.

   It is important to set the primary wavelength to measure peptide bonds. The peak absorbance (188 nm) cannot be used, however, because many of the HPLC reagents also absorb strongly in this region and obscure the signal from the protein. At 215 nm, the solvents do not interfere with the protein.

   The secondary tracings cover the absorption of aromatic amino acids. Although having

the whole chromatogram at these secondary wavelengths is not essential for this method, it can be useful (see Anticipated Results).

3. Set reference wavelength at 550 nm with a bandwidth of 100 nm.

   The reference wavelength is generally set outside the range of protein absorption. Because proteins do not absorb at this wavelength, a wide bandwidth is recommended to avoid spurious readings. The absorbance of the reference wavelength is subtracted from the absorbance of the sample wavelength to give the absorbance of the signal for the protein.

4. Determine the points at which diode-array spectra will be taken and saved. Set acquisition and storage of spectrum to "peak controlled."

   This means that diode-array spectra will be taken at a point on the upslope, at the apex, and at a point on the downslope of a peak. The information obtained will be used to assess peak purity. If a peak has shoulders or is asymmetrical, spectra will automatically be obtained at additional points using the peak purity program.

5. Set threshold to define height of signal expected for the smallest peak from 0.1 to 999 milliabsorbance units (mAU).

   A threshold of 5 mAU is suggested for a sample of 50 pmol protein.

   This parameter is relevant to the size of the record that will be made and saved during the run. Diode-array spectra will be saved only for peaks that are higher than the threshold height defined, not for peaks that are lower.

6. Set peakwidth to equal the half-height width expected for the narrowest peak, with limits from 0.007 to 10 min.

   A suggested starting value for this parameter is 0.15 min. Together with the threshold, the peakwidth determines the size of the record made for the run. Diode-array spectra will be saved only for peaks displaying a wider half-height width than the value chosen, not for peaks narrower than this width. This eliminates saving spectra for peaks that are due to noise.

7. Set stop time for acquisition of spectra.

   This value controls the diode-array detector only; it can be set to stop independently from the run, earlier than the run, or automatically with the run.

8. Set a gradient rate of 1% buffer B/min and a flow rate of 1 ml/min for the 4.6-mm-i.d. column, 200 µl/min for the 2.1-mm-i.d. column. Set the amount of buffer A plus the amount of buffer B to equal 100%.

   HPLC pumps are controlled by the B pump, which can vary from 0% to 100%. The amount of buffer A plus the amount of buffer B must equal 100%. Thus 0% buffer B means that the B pump is off and only the A pump is working (e.g., 100% A), and 100% buffer B means that the A pump is off and only the B pump is working. To achieve the same chromatographic separation on a 2.1-mm-i.d. column as obtained with a 4.6-mm-i.d. column, adjust the flow rate according to the ratio of the cross-sectional areas of the two columns. A 4.6-mm-i.d. column is roughly five times the cross-sectional area (πr^2) of a 2.1-mm-i.d. column.

   Gradient rate defines the change in the gradient over time. The gradient should not be started until the sample has had time to flow onto the column. For a sample 100 µl in volume and a flow rate of 1 ml/min, it will take 0.1 min to load the sample onto the column. The following table illustrates typical gradient conditions for a 100-µl sample and a flow rate of 1 ml/min. This information would be entered into a liquid chromatography (LC) time table. The time required for the run should be determined and

BCU LIBRARY CENTRAL UNIVERSITY LIBRARY

entered at the appropriate place in the program. This time limit should be long enough to allow for all of the events specified in the LC time table.

Time (min)	% B
0	10
1	10
51	60
56	80
66	10

After a trial run, if peaks of interest are too close to each other, the gradient rate can be made shallower (e.g., 0.5% buffer B/min) to spread out the gradient or steeper (e.g., 1.5% buffer B/min) to sharpen well-resolved peaks.

9. Equilibrate the column with 10% buffer B by running a volume equivalent to at least 3 column volumes.

10. Prepare experimental and control samples in 10% buffer B, keeping the volume to a minimum and making all samples the same volume.

A good control sample will contain everything present in the experimental sample except the protein. The following table illustrates a sample preparation:

	Control	Experimental
Protein sample	—	90 µl
Sample buffer	90 µl	—
HPLC buffer B	10 µl	10 µl

Quantities of protein in the sample can range from 0.5 to 5 nmol for a 4.6-mm-i.d. column and from 100 to 500 pmol for a 2.1-mm-i.d. column. For reversed phase chromatography, sample volumes of 20 to 100 µl will work.

11. Filter samples through a low-protein-absorption 0.21-µm pore size filter or microcentrifuge 5 min at maximum speed to remove particulates.

12. Choose an appropriate label for the control run. Inject the control sample into the HPLC apparatus.

The control sample is used to control for spectral impurities in buffers. The control should be run before the experimental sample to permit careful inspection of the entire gradient. The HPLC starts acquiring data on injection. If a rheodyne injector is used, leave the sample loop on line long enough to allow the whole sample to be applied to the column.

13. Choose an appropriate label for the experimental run. Inject the experimental sample into the HPLC.

As above, the injection of the sample starts the acquisition of data. The data is stored in a computer file under the given label.

Data analysis

14. Subtract the UV spectrum of the control from that for the protein to take into account impurities found in the sample buffer.

15. Run the peak purity program, which compares spectra for points at the apex and on the upslope and downslope of each peak.

16. Compare normalized overlaid spectra.

PGIMER-IASLC CENTRAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Acid/ethanol cleaning solution (3 N HCl/50% ethanol)

Add 50 ml of 12 N HCl to 50 ml deionized water. Add the resulting solution to 100 ml ethanol. Store several months at room temperature.

CAUTION: *Never pour water or ethanol into the concentrated acid. Always pour the acid into water or ethanol. Wear appropriate personal protection.*

Holmium oxide reference solution

4% holmium oxide in aqueous solution of 10% (v/v) perchloric acid sealed in a nonfluorescent fused silica cuvette of optical quality (National Institute of Standards and Technology, Standard Reference Materials catalog; see *SUPPLIERS APPENDIX*).

Store at room temperature for several months.

COMMENTARY

Background Information

Ultraviolet absorption spectroscopy played a key role in early investigations of protein structural transitions (Wetlaufer, 1962). The explosive growth of molecular biology and expanded use of more powerful spectroscopic techniques such as nuclear magnetic resonance (NMR), X-ray crystallography, circular dichroism, and fluorescence relegated this technique to use in simple assays and concentration determinations. During this period, however, electronic and optical systems improved considerably, allowing extraction of much more detailed information from simple near-UV spectra. It was soon shown that the magnitude of the near-UV second derivative peaks can be used for quantitative determination of the content of aromatic amino acids in proteins (Ichikawa and Terada, 1977; Servillo et al., 1982, Levine and Federici, 1982; Nozaki, 1990). These methods employ either peak-to-peak distances or least-squares multicomponent analysis and are capable of resolving all three components simultaneously. In fact, the spectrum of a mixture of three proteins can be easily resolved by this method if the spectra of the individual proteins are known (Mach et al., 1989). The composition of simple protein mixtures and complexes can then be monitored without actual chromatographic separation.

Generally, the spectra of the individual aromatic amino acids become red-shifted (move to higher wavelengths) upon incorporation into protein structures (see Fig. 7.2.1). These shifts can be as large as 5 nm for Tyr, 4 nm for Trp, and 2 nm for Phe residues (Mach and Middaugh, 1994). Of course, many side chains naturally exist in quite polar microenvironments in native structures, so these shifts are not always so large. Unfolding of a protein usually results in some increased solvent exposure of aromatic side chains and, consequently, in blue shifts (toward lower wavelengths). If the spectral signal near 295 nm (Fig. 7.2.2) shifts left even a fraction of a nanometer, the absorbance value at any given wavelength in this vicinity may change significantly because of the steepness of the slope. This phenomenon can be effectively utilized to follow the unfolding of a protein by measuring the spectral intensity at this wavelength as a function of a structure-perturbing variable (e.g., temperature, pH, protein concentration, or urea concentration) or by measuring the difference spectra between two proteins of identical concentration, one of which is subject to structural perturbation. The resulting values, however, are relative. Use of the intersection of a portion of the second-derivative spectrum with the wavelength axis is therefore recommended.

Application of more complex fitting routines and use of systematically shifted standard spectra allow the individual spectra of Trp, Tyr, and Phe to be resolved from the composite protein spectrum (Mach and Middaugh, 1994). This allows the fractional exposure of aromatic residues to be assessed using additions of apolar compounds as perturbants. The observed shifts are compared to the shifts in model, solvent-exposed compounds (e.g., acetylated and amidated aromatic amino acid derivatives), and the solvent exposure is estimated from a comparison of these values. In addition to proteins, DNA and RNA possess strongly absorbing chromophores (purine and pyrimidine bases) in the near-UV (Table 7.2.1), result-

BRANDEIS UNIVERSITY LIBRARY

ing in the well-known featureless spectrum of nucleic acids with a peak near 260 nm. The molar absorptivity (extinction coefficient) of polynucleotides is strongly structure dependent and this is often utilized in monitoring thermal structural transitions (i.e., in melting experiments). In addition, DNA and RNA have well-defined, albeit weak, second-derivative spectra that can be employed to determine the amount of protein in the presence of DNA (Mach et al., 1992a).

Very strong absorbance is also produced by all proteins in the far-UV range (180 to 250 nm). This is primarily the result of the intense absorption by the peptide bond in the far-UV. Because of amide absorption near 188 to 190 nm, a peak is actually present, but the intense absorption by oxygen as well as most solutes in this region makes it difficult to resolve experimentally. In some spectrophotometers, this peak can be resolved by flushing the instrument with N_2 and by the use of UV-transparent anions (e.g., phosphate or perchlorate), but this is rarely done today. Although the structure and position of this peak contain secondary-structure information as well as side-chain contributions, the peak is rarely used experimentally because of the superiority of circular dichroism measurements in the same region. It is occasionally employed, however, for concentration measurements or simple protein detection during chromatography by monitoring wavelengths on the slope of the peak (e.g., 210 and 215 nm). If the protein or peptide studied lacks aromatic amino acids, or the concentration of the sample does not produce sufficient signal, its concentration can be calculated by employing a set of average extinction coefficient values in the far-UV range that were obtained based on the analysis of twelve proteins (Scopes, 1974):

(1) E(1 mg/ml, 1 cm) at 205 nm = 31 with stated error of 5%

(2) E(1 mg/ml, 1 cm) at 205 nm = 27.0 + 120 A_{280}/A_{205} with stated error of 2%.

Extreme care should be taken when correcting for the many buffer components absorbing in this region.

In principle, methods such as SDS-polyacrylamide gel electrophoresis (SDS-PAGE; *UNITS 10.1 & 10.3*) are more effective than UV spectroscopy in assessing the purity of a protein sample, based on detection of the presence and absence of additional bands. Such methods, however, may fail to ascertain contaminating proteins of similar molecular weight, unless isoelectric focusing (IEF) gels are used (*UNITS 10.2 & 10.8*). In contrast to electrophoretic (Chapter 10) or chemical (Chapter 11) analyses, the UV absorption technique is both nondestructive and rapid. Moreover, only a modest amount of sample is required.

This unit describes recombinant protein analyses using the two general types of UV spectrophotometers (Fig. 7.2.3): conventional scanning instruments, in which light absorption is measured by mechanically moving a monochromator allowing data to be accumulated temporally as a function of wavelength (Basic Protocol 1 and Support Protocol) and diode-array spectrophotometers (Basic Protocol 2), in which a light beam containing all frequencies is simultaneously split onto an array of detectors (diodes) after passing through the sample. Diode-array spectrophotometers (Hewlett-Packard 8450 to 8453 series or Beckman DU 7000) are simpler and consequently have excellent wavelength reproducibility. A single spectrum is typically acquired in 0.1 sec; longer acquisition times provide an average of a number of measurements. Spectral bandwidth is fixed at 1 or 2 nm. The small cross-sectional area of the light beam (~1×5 mm for the Hewlett-Packard diode-array instrument) allows small sample volumes and free choice of cuvette format. Such instruments can be easily used on-line to collect and analyze spectra of protein fractions leaving a chromatographic column—e.g., from gel filtration (*UNIT 8.3*), ion-exchange (*UNIT 8.2*), and reversed-phase media. In contrast, conventional scanning spectrophotometers require longer acquisition times, since at any given time only one wavelength is read.

Critical Parameters

The spectrophotometers contain moving internal parts that require frequent wavelength calibration. Of great importance when using scanning spectrophotometers is the setting of proper detector response time relative to scanning speed. A decrease in absorbance values, broadening of bands, and shifts are often observed in such cases.

Although the intrinsic absorbance of the aromatic amino acids varies little with temperature, it is advisable in many cases to control the temperature of the sample. Many processes destructive to proteins (e.g., oxidation, deamidation, unfolding, and aggregation) are accelerated at elevated temperature (Volkin and Middaugh, 1992). Under most conditions of humidity, however, a temperature of 15°C should be regarded as minimal. Decreasing the

BOU IAS CENTRE UNIVERSITY LIBRARY

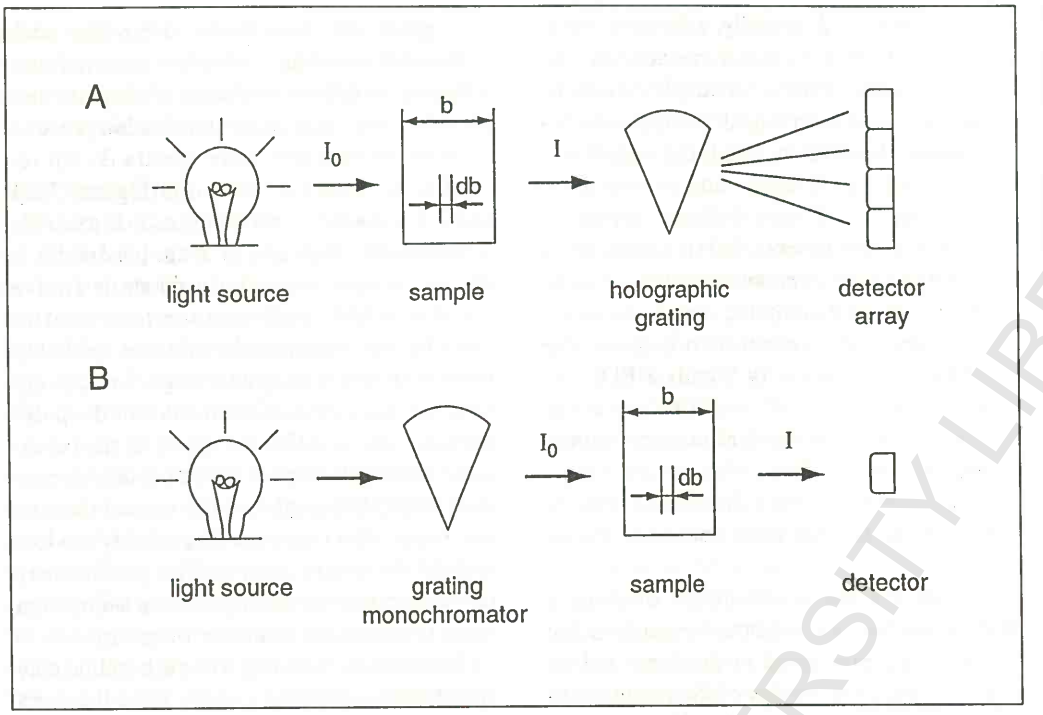


Figure 7.2.3 Basic scheme for spectrophotometric measurements: lamp, sample and detector; b (path) and db are shown in the sample and *I* and I_0 are denoted (see below). **(A)** Diode-array and **(B)** conventional scanning spectrophotometers. The relative decrease of the intensity of the light beam after passing through the sample is proportional to the distance (db) and the concentration of the absorbing component: $-db/b = kc \times db$, where k is an absorption coefficient. Upon integration, the result is $\ln(I/I_0) = -kcl$. The fraction of light transmitted through the sample is known as the transmittance $T = I/I_0$. Absorbance is defined as $A = \log I_0/I = \varepsilon cb$, where ε is a molar absorptivity (extinction coefficient), c is the molar concentration, and b is the path length (in centimeters). Alternatively, weight extinction coefficients $E_{1\ cm,\ 0.1\%}$ expressed per 1 mg/ml (0.1%) concentration are used.

temperature below this level may result in water condensing on the optical surface of the cuvette. The power of deuterium lamps used in UV spectrophotometers is usually not sufficient to cause photooxidation of aromatic residues. However, if the fluorescence spectra were acquired using the same sample, some of the tryptophan residues may have been oxidized to *N*-formylkynurenine and kynurenine by light from more powerful xenon lamps. This condition can be recognized by the appearance of a broad spectral band extending over 350 nm (similar to light scattering) and can be identified by comparing spectra as a function of irradiation time in a fluorometer or by direct observation of photoproduct fluorescence bands at ~400 nm (with excitation at 320 nm). Solvent-accessible tryptophan side chains are particularly prone to photooxidation, with half-lives as short as several minutes (Mach et al., 1995).

An aqueous buffer in which the protein of interest is stable should always be used. For example, phosphate-buffered saline, pH 7.4 (PBS; *APPENDIX 2E*), or an equivalent buffer containing 10 mM sodium phosphate and 120 to 150 mM sodium chloride, pH 7.0 to 7.4, is often employed. Choice of pH is important, because the spectral contribution of absorbing residues may be directly or indirectly (effect on protein conformation) pH dependent. Whereas tyrosine ionization (pK=10) results in an entirely different spectrum (λmax = 294 nm), ionization of nonabsorbing or weakly absorbing side chains (e.g., lysine, pK~10; aspartic and glutamic acids, pK~4.5; histidine, pK~6.5; or cysteine, pK~8.5) may produce subtle shifts if the microenvironments of UV-absorbing chromophores or the overall conformation are altered. The pH should be sufficiently different from the isoelectric point to ensure good solubility. Note that many agents used to improve protein solubility (e.g., surfactants, denaturing agents, and sugars; Middaugh and Volkin, 1992) may affect spectral band positions.

For HPLC analysis, it is important that HPLC-grade solvents, free of UV-absorbing contaminants, be used. Similarly, the sample

PRINCETON UNIVERSITY LIBRARY

should be prepared carefully, with the highest-grade components, to avoid inadvertently introducing contaminants. All samples should be either filtered or centrifuged to remove particulate matter. In addition, HPLC tracings should be examined for shoulders and unusual peak broadening, which may indicate impurities (Zavitsanos and Goetz, 1991). Diode-array spectra will be obtained automatically at additional points if the computer detects an asymmetrical peak. This information is invaluable in determining peak purity. Finally, HPLC column performance should be routinely assessed by injecting a set of standard proteins or peptides (Mant and Hodges, 1991), as poor resolution may broaden peaks. Resolution deteriorates on columns that have been used excessively.

One of the major advantages of using a diode-array detector to determine purity is that it is relatively unaffected by gradients and requires no prior knowledge of the sample. Absence of a strong absorption band between 200 and 230 nm indicates that the analyzed peak is not a protein but rather a nonproteinaceous contaminant, such as a detergent or a chromophore, picked up during the purification procedure.

Troubleshooting

A few recurring problems are encountered in the analysis of proteins by UV spectrophotometry. A major, often unrecognized difficulty is the phenomenon known as absorption flattening. Imagine a large number of protein molecules aggregating so that some incident light is able to pass "in between" them without encountering an absorbing entity. However, because of the high concentration of chromophores in the aggregated particle, photons that happen to encounter the aggregate are unlikely to emerge on the other side. As a result, the intensity of such spectra are significantly reduced and often distorted. If this problem is suspected, the length of the effective light path should be changed. If absolute values of protein concentration are found to be pathway dependent, this phenomenon may be present.

The presence of anomalies such as significant curvature, negative values, or "bumps" in the 300- to 350-nm spectral region indicates either technical problems or the presence of unknown absorbing components, perhaps due to contaminating solutes. The spectra of any suspected absorbing impurities can be measured and compared to the experimental spectra to help determine if impurities are present. Detergents are often found to produce such extraneous absorbance, and their potential contribution should be evaluated if they are employed at any stage in the purification process.

If the second-derivative spectra do not resemble the examples shown in Figures 7.2.1 and 7.2.2, the most probable reason is too wide a bandwidth. Typically, a 2-nm bandwidth is the largest value permitting accurate derivative calculation. If the derivative spectrum is calculated by the instrument's software, probably more than five data points (e.g., 7 to 21) are used for calculation of each value or the polynomial order is different from 2. If the calculated second-derivative spectra resemble random noise, the signal-to-noise ratio of the normal (zero-order) spectrum is probably too low to allow this type of analysis. This problem may be overcome by increasing spectral acquisition times or averaging a number of spectra.

Problems in obtaining a stable baseline may arise from a number of factors. If the incorrect lamp was turned on—the tungsten-halogen (or "visible") instead of the deuterium (or "UV") lamp—spectra will not be obtainable. Failure of the UV lamp will usually manifest itself in an inability to obtain zero absorbance values during a "blank" measurement after collecting the baseline with both cuvettes containing the buffer. Many instruments contain an indicator of this condition. Care should be taken not to directly inspect the deuterium lamp for ignition, as UV irradiation can seriously damage the retina of the eye. The cuvette itself should also be reexamined. Only synthetic quartz provides sufficient transparency over the UV region. The inadvertent use of a glass cuvette will block passage of light through the system. If spectra cannot be obtained, the cuvette may have been positioned incorrectly, with opaque or black walls perpendicular to the light beam.

Instability of the absorbance signal can occur for a number of other reasons, such as liquid smears, water vapor condensation or fingerprints on the optical surfaces, or air bubbles or other suspended or sedimenting particles in the test solution. As indicated previously, the meniscus should be above the upper edge of the beam and no light should be able to bypass the sample. In some cases, the problem of insufficient volume of sample available can be solved by inserting some space-filling material beneath the cuvette.

In diode-array detection schemes, it is important to run a control sample consisting of the solution in which the protein is dissolved, to detect possible contaminants. Threshold and

BIBLIOTECA UNIVERSITARIA

peak width values should be set correctly to ensure that all peaks of interest are seen by the diode-array detector. Otherwise, small contaminating peaks may be skipped.

Anticipated Results

Figure 7.2.2A illustrates a typical example of a near-UV protein spectrum. Despite differences in aromatic amino acid compositions in proteins, spectra are usually qualitatively similar, with a maximum falling between 276 and 282 nm and only marginal absorbance above 300 nm. Small soluble monomeric proteins do not usually produce any light scattering. The second-derivative pattern should have the distinct features shown in Figure 7.2.2B characteristic of tryptophan, tyrosine, and phenylalanine residues, if these are present in the sequence. The position of the intersect with the wavelength axis of the declining second-derivative region near 290 nm is usually found at 289.41 ±0.61 nm using 1-nm data point spacing (based on the average from 14 tryptophan-containing proteins). This value should be ~0.5 nm lower when the 2-nm data point interval is used. The reproducibility should be within 0.05 nm for pure and structurally intact proteins.

The diode-array detector obtains spectra of very high quality and precision, enabling the operator to check the purity of a protein by comparing normalized spectra from different points of the eluted peak. It uses information within the protein (i.e., its amino acid composition) to check purity. Proteins of differing composition will have different spectral absorptions, particularly if they vary widely in aromatic amino acid content. Thus, if the retention times of two proteins are close enough, the proteins appear to be in the same peak. Based on their absorption characteristics the diode-array detector can distinguish them by comparing normalized spectra at the front end of the peak on the upward slope, at the back end of the peak on the downward slope, and at the apex. If the protein is pure, the spectra collected from these three points will superimpose exactly. If the protein is impure, the spectra collected from the three points will deviate, aside from the rare case in which the impurity and the sample exactly coelute from the column. The general rule is that small-to-moderate levels of impurity can be detected in a contaminated sample, provided the spectral properties of the contaminant vary from the recombinant protein.

The examples in Figures 7.2.4 and 7.2.5 illustrate pure and impure samples. Figures 7.2.4A and 7.2.5A show the main chromatography tracings at 215 nm. The first visible peak in each chromatogram is the injection peak and the second is the protein of interest. Figures 7.2.4B and 7.2.5B show the stored diode-array spectra for points collected on the upswing, the apex, and the downswing of the protein peak of interest. A pure protein is characterized by identical spectra at these points, whereas an impure protein will have different spectra at each point. In Figure 7.2.4B, the spectra are close to identical, indicating that the protein is fairly pure. In Figure 7.2.5B, on the other hand, the spectra vary widely. Diode-array spectra for points 1 and 2 were collected early in the peak. These spectra have little absorbance in the region from 200 to 220 nm, indicating that a nonproteinaceous component is eluting at the front end of the BSA peak. Peaks 3, 4, and 5 correspond to the upswing, the apex, and the downswing, respectively, of the main peak. The diode-array spectra of these points are different enough to indicate contamination. Figures 7.2.4C and 7.2.5C show the superposition of tracings measured at the primary, secondary, and tertiary wavelengths. Note that these closely superimpose for the pure sample (Fig. 7.2.4C) but not for the impure sample (Fig. 7.2.5C).

Time Considerations

The entire procedure for measurement and analysis of protein UV spectra can usually be completed within 30 min, of which 20 min is needed for lamp warm-up and sample preparation. This time should be shorter if the instrument is already turned on. Diode-array spectrophotometers interfaced to computers equipped with data processing programs are especially time efficient when a larger number of samples is analyzed. Characteristically, even the most complex analysis routines, when implemented within the existing software, usually require only seconds for completion. In the case of conventional scanning instruments, calibration (measurements of standards) combined with slower scanning rates may require some additional time.

An HPLC run takes ~2 hr, including the column equilibration with buffer A. Examination of the diode-array detector information should take at most 1 additional hour for the novice and 30 min for the experienced operator.

BOYNES CENTRAL LIBRARY

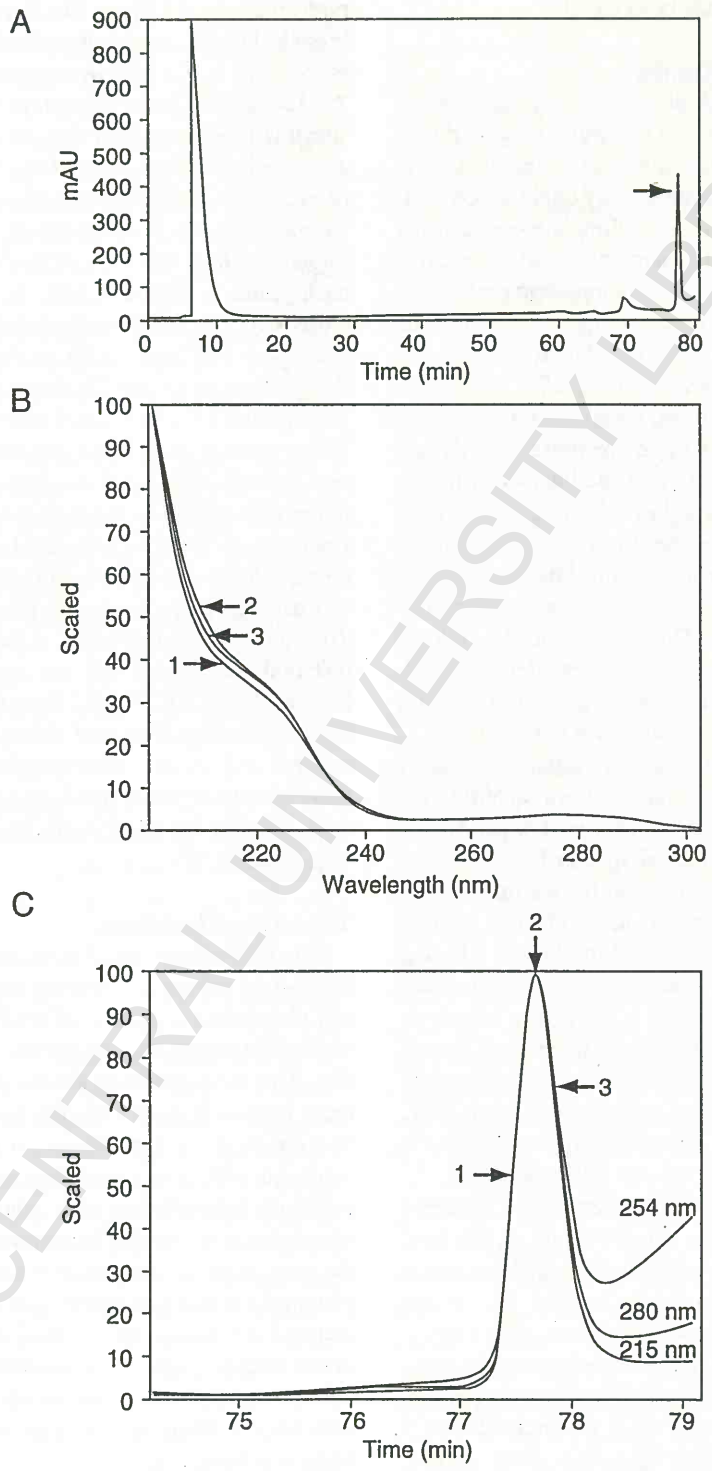


Figure 7.2.4 Spectra of a pure protein sample (carbonic anhydrase). **(A)** Main chromatography tracings at 215 nm. Arrow indicates retention time of the component analyzed. **(B)** Stored diode-array spectra for points collected on the upswing, the apex, and the downswing of the protein peak. **(C)** Superposition of tracings measured at the primary, secondary, and tertiary wavelengths.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

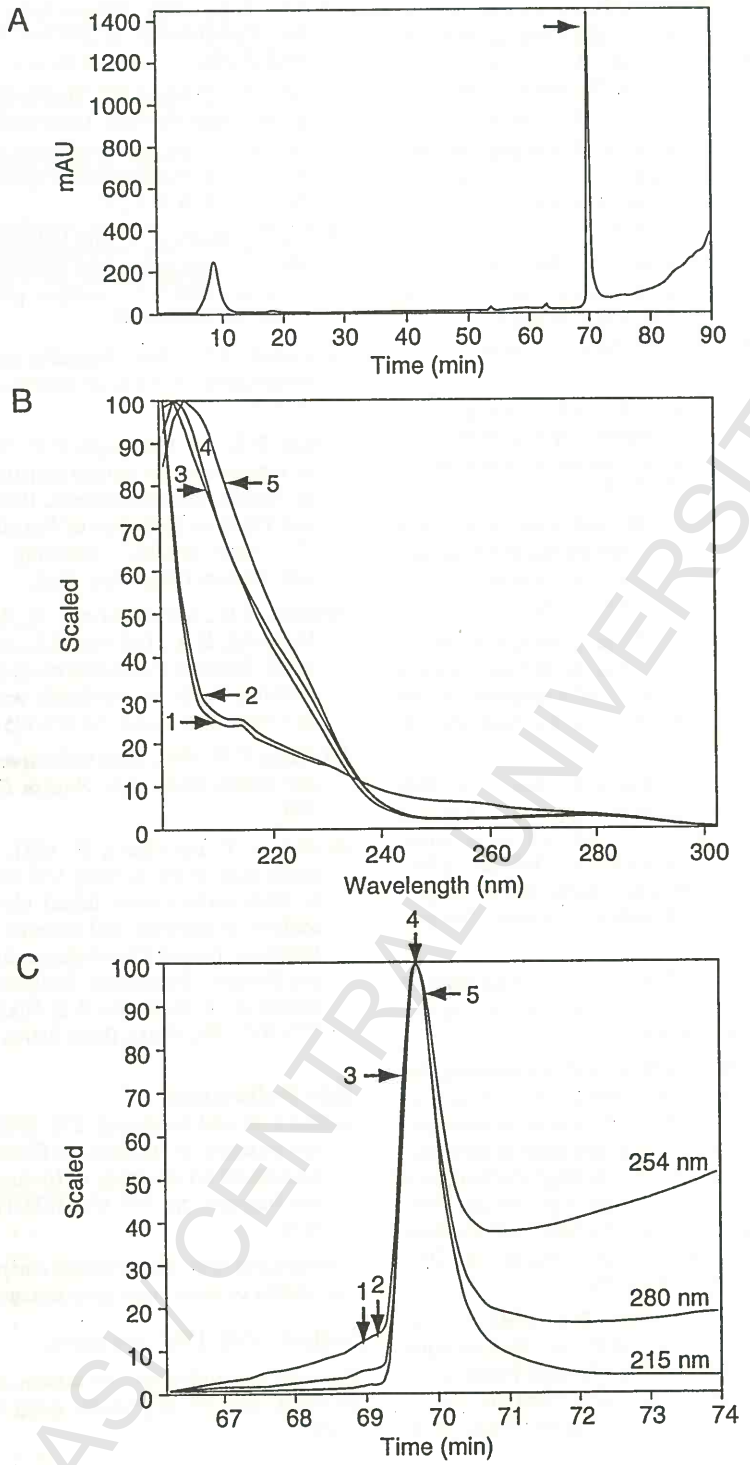


Figure 7.2.5 Spectra of an impure protein sample (bovine serum albumin). **(A)** Main chromatography tracings at 215 nm. Arrow indicates retention time of the component analyzed. **(B)** Stored diode-array spectra for points collected on the upswing, the apex, and the downswing of the protein peak. **(C)** Superposition of tracings measured at the primary, secondary, and tertiary wavelengths.

BRAU IIAS / CENTRAL UNIVERSITY LIBRARY

Literature Cited

Glasel, J.A. 1995. Validity of nucleic acid purities monitored by 260 nm/280 nm absorbance ratios. *BioTechniques* 18:62-63.

Ichikawa, T. and Terada, H. 1977. Second derivative spectrophotometry as an effective tool for examining phenylalanine residues in proteins. *Biochim. Biophys. Acta* 494:267-270.

Levine, R.L. and Federici, M.M. 1982. Quantitation of aromatic residues in proteins: Model compounds for second-derivative spectroscopy. *Biochemistry* 21:2600-2606.

Mach, H. and Middaugh, C.R. 1994. Simultaneous monitoring of the environment of tryptophan, tyrosine and phenylalanine residues in proteins by near-UV second derivative spectroscopy. *Anal. Biochem.* 222:323-331.

Mach, H., Thomson, J.A., and Middaugh, C.R. 1989. Quantitative analysis of protein mixtures by second derivative absorption spectroscopy. *Anal. Biochem.* 181:79-85.

Mach, H., Middaugh, C.R., and Lewis, R.V. 1992a. Detection of proteins and phenol in DNA samples with second derivative absorption spectroscopy. *Anal. Biochem.* 200:20-26.

Mach, H., Middaugh, C.R., and Lewis, R.V. 1992b. Statistical determination of the average values of the extinction coeficients of tryptophan and tyrosine in native proteins. *Anal. Biochem.* 200:74-80.

Mach, H., Burke, C.J., Sanyal, G., Tsai, P.-K., Volkin, D.B., and Middaugh, C.R. 1995. Origin of the photosensitivity of a monoclonal immunoglobulin G. *In* Formulation and Delivery of Proteins and Peptides (J.L. Cleland and R. Langer, eds.) pp. 72-84. American Chemical Society, Washington, D.C.

Manchester, K.L. 1995. Value of A_{260}/A_{280} ratios of measurement of purity of nucleic acids. *BioTechniques* 19:208-210.

Mant, C.T. and Hodges, R.S. 1991. Requirement for peptide standards to monitor column performance and the effect of column dimensions, organic modifiers, and temperature in reversed-phase chromatography. *In* High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation (C.T. Mant and R.S. Hodges, eds.) pp. 289-295. CRC Press, Boca Raton, Fla.

Middaugh, C.R. and Volkin, D.B. 1992. Protein solubility. *In* Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation (T.J. Ahern and M.C. Manning, eds.) pp. 109-134. Plenum Press, New York.

Nozaki, Y. 1990. Determination of tryptophan, tyrosine and phenylalanine by second derivative spectrophotometry. *Arch. Biochem. Biophys.* 277:324-333.

Savitzky, A. and Golay, M.J.E. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36:1627-1639.

Scopes, R.K. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* 59:277-282.

Servillo, L., Colonna, G., Balestrieri, C., Ragone, R., and Irace, G. 1982. Simultaneous determination of tyrosine and tryptophan residues in proteins by second-derivative spectroscopy. *Anal. Biochem.* 126:251-257.

Steiner, J., Termonia, Y., and Deltour, J. 1972. Comments on smoothing and differentiation of data by simplified least squares procedure. *Anal. Chem.* 44:1906-1909.

Timasheff, S.N. 1966. Turbidity as a criterion of coagulation. *J. Colloid Interface Sci.* 21:489-497.

Volkin, D.B. and Middaugh, C.R. 1992. The effect of temperature on protein structure. *In* Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation (T.J. Ahern and M.C. Manning, eds.) pp. 215-247. Plenum Press, New York.

Weidner, V.R., Mavrodineanu, R., Mielenz, K.D., Velapoldi, R.A., Eckerle, K.L., and Adams, B. 1985. Spectral transmittance characteristics of holmium oxide in perchloric acid solution. *J. Res. Natl. Bur. Stand.* 90:115-125.

Wetlaufer, D.B. 1962. Ultraviolet spectra of proteins and amino acids. *Adv. Protein Chem.* 17:303-390.

Zavitsanos, P. and Goetz, H. 1991. The practical application of diode array UV-visible detectors to high-performance liquid chromatography analysis of peptides and proteins. *In* High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation (C.T. Mant and R.S. Hodges, eds.) pp. 553-562. CRC Press, Boca Raton, Fla.

Key References

Cantor, C.R. and Schimmel, P.R. 1980. Absorption spectroscopy. *In* Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function, pp. 349-408. W.H. Freeman, New York.

Provides theoretical background and practical considerations in absorption spectroscopy.

Wetlaufer, D.B. 1962. See above.

Provides comprehensive discussion of factors involved in analysis of proteins using UV spectroscopy.

Savitzky, A. and Golay, M.J.E. 1964. See above.

Provides formulas for smoothing and differentiation of data.

BCU IASI JOURNAL UNIVERSITY LIBRARY

Steiner, J. et al. 1974. See above.

Provides formulas for smoothing and differentiation of data.

Levine, R.L. and Federici, M.M. 1982. See above.

Describes an application of second-derivative spectroscopy to analyze aromatic amino acid content employing a diode-array spectrophotometer.

Mach, H., Middaugh, C.R., and Lewis, R.V. 1992b. See above.

Describes calculation of protein extinction coefficients based on sequence and provides a formula for light-scattering correction at 280 nm.

Contributed by Henryk Mach and C. Russell Middaugh
Merck Research Laboratories
West Point, Pennsylvania

Nancy Denslow
University of Florida
Gainesville, Florida

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Determining the Identity and Structure of Recombinant Proteins

After isolating a recombinant-derived protein and establishing that it is pure, it is necessary to confirm that it has the same primary sequence as predicted from the cDNA coding sequence. There are several reasons why this may not be the case: (1) errors in sequencing the gene (e.g., single-base errors resulting in single amino acid substitutions) and typographical errors in translating the sequence; (2) translation errors leading to the misincorporation of various amino acids (e.g., Harris et al., 1993; Randhawa et al., 1994); (3) unexpected translation start and finish points (see *UNIT 6.1* for sample references); (4) retention or nonretention of N-terminal methionine when expressing proteins in *E. coli* (*UNIT 6.1* and *UNIT 7.1*); (5) proteolytic processing by host cell proteases or, if the expressed protein is a protease, autolysis; and (6) chemical modification of specific amino acids: e.g., deamidation of amides (Asn and Gln), oxidation of cysteine and methionine, and, if urea was used for protein purification or folding, carbamylation of lysine residues (*UNIT 6.1*). Although this list is not exhaustive, it is enough to indicate that there are sufficient reasons to warrant a careful analysis of the primary sequence of the isolated protein to confirm that it corresponds to that expected from the cDNA sequence.

The standard methods for confirming the primary sequence of a recombinant protein are combinations of (1) accurate amino acid analysis (*UNIT 3.2*) and mass spectrometry (Stults, 1995); (2) N- and C-terminal amino acid sequence determination by Edman degradation; and (3) peptide mapping in conjunction with the first two methodologies.

In this unit peptide mapping protocols with separation of the constituent peptides by high-performance liquid chromatography (HPLC) analysis (Basic Protocol 1) and by high-resolution SDS-polyacrylamide gel electrophoresis (Alternate Protocol) are presented. The latter approach is included for the benefit of laboratories that do not have access to the more expensive HPLC equipment. Other peptide mapping techniques are described in *UNIT 11.6*.

When peptide mapping a recombinant protein, the peptide bonds are first cleaved using specific proteases or chemical reagents to produce a mixture of peptides. The composition of the mixture is a function of both the primary sequence and the specificity of the protease or chemical reagent. In the protocols described, proteases with strict specificity are used, and as the translated protein sequence is known, it is possible to predict the number and size of the peptides. The composition of the peptides can be determined by mass spectrometry and/or amino acid analysis (*UNIT 3.2*). Peptide mapping is ideally suited for comparative purposes—for example, combined analysis of the recombinant protein and its natural counterpart (or some other well-characterized standard). The method also has the advantage of reviewing the whole sequence.

The most common post-translational modification, regardless of whether the protein is derived from a prokaryotic or eukaryotic host, is the oxidation of paired cysteine residues to form disulfide bonds. Proteins can contain one or more intramolecular disulfide bonds (for a general review, see Thornton, 1981). For dimeric proteins, intermolecular disulfide bonds can occur between the subunits, as in interleukin-5 (Proudfoot et al., 1991) and macrophage colony-stimulating factor (Glocker, et al., 1993). Some dimeric proteins contain both intra- and intermolecular disulfide bonds; the best known example is insulin (Sanger, 1959).

In this unit the general strategy used to determine the linkage pattern of a monomeric recombinant protein containing two intramolecular disulfide bonds is described (Basic

Copyright © 1996 by John Wiley & Sons, Inc.

POLYTECHNIC UNIVERSITY LIBRARY

Protocol 2). The approach used is an extension of peptide mapping, where the aim is to isolate and characterize peptides containing only a single disulfide bond.

In addition to the standard protein chemical approaches, isoelectric focusing in polyacrylamide gels can indicate charge heterogeneity arising from specific and nonspecific post-translational modifications. The usefulness of isoelectric focusing is increased when it is combined with electrophoresis. The resultant two-dimensional electrophoretic method, in which the protein isoelectric point is displayed as a function of pH, is described as electrophoretic titration curve analysis (Basic Protocol 3). This method is especially useful for checking for deamidation (e.g., of Asn to Asp) in which additional negative charge is introduced into the modified protein (Wingfield et al., 1987). Other modifications that can introduce charge heterogeneity include oxidation of cysteine residues, carbamylation of lysine side chains resulting from use of urea, and internal proteolytic clipping in which additional N- and C-termini are introduced (i.e., within a disulfide loop; Niekamp et al., 1969). Proteins expressed in eukaryotic systems can be subject to specific post-translational modification such as glycosylation and phosphorylation (Chapters 12 and 13, respectively), which can also introduce charge heterogeneity. Proteins produced in *E. coli* can also be subject to certain post-translational modifications such as acetylation of the N terminus and of the ε-side chain of lysine. These modifications will also result in charge heterogeneity (Violand et al., 1994).

Electrophoretic titration curve analysis is also extremely useful for predicting suitable conditions for ion-exchange chromatography (see *UNIT 6.1* for further details).

BASIC PROTOCOL 1

PEPTIDE MAPPING BY HPLC

The most important way to characterize a recombinant protein is to show that its primary amino acid sequence matches the predicted translation of the DNA coding sequence. In theory, one can simply digest the protein with an enzyme or chemical agent that is highly specific and compare the resulting fragments with those expected. There are computer programs that predict fragmentation patterns for proteins with all of the commonly used enzymes or chemical agents. In practice, however, the comparison can be a little tricky, because often digestions do not go to completion or peptide fragments behave differently than predicted due to slight variations in technique.

This section covers protocols for two different strategies for peptide mapping. The first, based on high-performance liquid chromatography (HPLC), provides the best resolution of the digested protein and can be used in tandem with other analytical techniques to identify each fragment if this is necessary. The second method is based on high-resolution SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Though less powerful than HPLC, in the sense of providing lower resolution of fragments, SDS-PAGE offers the advantage of allowing comparison to minute amounts of native protein, by using either very sensitive stains or radiolabeled protein. SDS-PAGE is also more feasible than HPLC in some laboratories.

Peptide mapping by HPLC is the method of choice if native material is not available for comparison. In this event, HPLC fractions must be collected for further analyses, including amino acid compositional analysis and N-terminal sequencing. New methods, recently developed, that couple peptide mapping by HPLC with mass spectrometry, are accurate enough to determine to within 0.01% the masses of the components of a digest of a recombinant protein. Tandem MS methods can even provide extensive sequence information. However, mass spectrometers may not be easily accessible and are not absolutely required for this analysis.

BAU IASI CENTRAL UNIVERSITY LIBRARY

Of course, the quickest way to determine the primary structure of a recombinant protein is to compare it to the native material. In this case, both the native material and the recombinant protein are digested under similar conditions and the resulting chromatograms (or SDS gels) are simply compared. Often identical chromatograms (or patterns in a gel) suffice to identify the recombinant protein in a preliminary way. However, native material is not always available, and when it is, it may be available only in minute quantities.

For all of these procedures, a recombinant protein is first denatured and alkylated and is then digested with a specific protease or chemical agent (in these protocols endoproteinase Lys-C, which specifically cleaves at Lys residues, is used as an example). The resulting fragments are analyzed by HPLC (Basic Protocol 1) or by high-resolution SDS-PAGE (Alternate Protocol). Another option presented that may be convenient in some cases is to perform the protein digestion or chemical cleavage in an SDS-polyacrylamide gel slice (see Support Protocol). Of course, both methods can be readily adapted to a number of other specific proteases or chemical cleaving agents. These are discussed in the Commentary section; specific protocols are given in *UNITS 11.1 & 11.4*.

Materials

Samples of recombinant and, if possible, native protein
Acetone/1 M HCl, –20°C (acidified acetone; optional)
8 M urea/50 mM Tris·Cl, pH 7.5
45 mM dithiothreitol (DTT)
100 mM iodoacetamide, prepared fresh and kept wrapped in foil
Sequencing-grade protease (*UNIT 11.1*) or chemical cleavage agent (*UNIT 11.4*):
e.g., endoproteinase Lys-C (Wako Chemicals)
Digestion buffer (see recipe)
HPLC solvents A and B for peptide mapping (see recipes)

50°C water bath or heating block
15- to 25-cm, 2.5-mm-i.d. C18 HPLC column (e.g., Vydac C18, 300 Å, 5-µm bead)
Analytical high-performance liquid chromatography system

Additional reagents and equipment for proteolytic digestion (*UNIT 11.1*) and/or chemical cleavage (*UNIT 11.4*) of proteins in solution and for digestion in SDS-polyacrylamide gel (optional; see Support Protocol)

Prepare samples and controls

1. Prepare a 50- to 300-pmol protein sample and, if available, an equal amount of the native protein (as control) by lyophilizing or by precipitating from solution with 9 vol of –20°C acidified acetone.

   In experiments in which both the recombinant protein and the native protein are used, the amounts of both should be similar; refer to UNITS 11.1 & 11.4 for appropriate protocols. Proteolytic digests can also be done directly in a gel slice for proteins purified by gel electrophoresis; in this case steps 1 to 6 of this protocol should be replaced by gel purification and digestion in the gel slice (see Support Protocol and UNIT 11.3).

   In the event that the protein cannot be precipitated or lyophilized, the conditions can be adjusted by adding solid urea directly to the protein sample, then adding DTT to the appropriate concentration and incubating (see step 3), before continuing with alkylation (step 4). In this case, expect that solid urea added to 8 M concentration will double the volume of the sample.

2. Label three microcentrifuge tubes: one for the recombinant protein, one for a control consisting of enzyme (or chemical cleavage agent) alone, and one for the native protein, if available.

BOSTON UNIVERSITY MEDICAL LIBRARY

Denature and alkylate protein

3. Place each sample in the appropriately labeled microcentrifuge tube, add 50 µl of 8 M urea/50 mM Tris·Cl, pH 7.5, and dissolve. Add 5 µl of 45 mM dithiothreitol. Incubate 15 min at 50°C.

   In the presence of 8 M urea and dithiothreitol, the protein sample becomes completely denatured and all disulfide bonds are reduced.

4. Add 5 µl of 100 mM iodoacetamide to alkylate cysteine residues. Incubate 15 min at room temperature in the dark.

   Iodoacetamide solution should be made fresh and kept protected from light. Wrapping the tube in foil works well. If a larger volume is used in step 3, the volume of iodoacetamide should be adjusted proportionately. The iodoacetamide concentration should always be at least twice the concentration of sulfhydryls, including those present in the protein and in the reducing agent, DTT, added to reduce disulfides.

Digest or chemically cleave the sample

5a. *For enzymatic digestion:* Add an appropriate quantity of the selected protease in digestion buffer: e.g., 140 µl digestion buffer containing 0.003 U endoproteinase Lys-C per µg protein, or ~1:25 (w/w) enzyme/protein. Cap tube, wrap a strip of Parafilm tightly around the cap to prevent evaporation, and allow to digest overnight (~16 hr) at 37°C.

   Refer to UNIT 11.1 for protocols for different proteolytic enzymes, including appropriate digestion buffers.

   It is important to reduce the urea concentration to ≤2 M for this step. Proteolytic enzymes are not active in higher urea concentrations because they become denatured. Most proteolytic agents will tolerate up to 10% acetonitrile as well; several published procedures include 10% acetonitrile in the digestion buffer (Fernandez et al., 1992).

5b. *For chemical cleavage*: Precipitate the reduced and alkylated protein with 9 vol ice-cold (–20°C) acidified acetone for 2 hr in a dry ice/ethanol bath and microcentrifuge 15 min at maximum speed, 4°C, to collect the precipitate. Solubilize the sample in the solvent appropriate for the cleavage.

   Specific conditions for use of CNBr, which cleaves at Met groups; BNPS-skatole, which cleaves at Trp groups; or hydroxylamine, which cleaves at Asn-Gly bonds, can be found in UNIT 11.4.

6. Microcentrifuge 15 min at maximum speed, room temperature, to pellet any particulate material that might have formed during the digestion or cleavage.

   The sample is now suitable for direct injection onto the HPLC column.

Perform HPLC analysis

7. Equilibrate a C18 column with HPLC solvent A, setting flow rate at 150 µl/min for a 2.1-mm-i.d. column.

   If a column with a wider i.d. is used (e.g., 4.6 mm), the flow rate should be adjusted in proportion to the increase in the cross-sectional area of the column. A 4.6-mm-i.d. column has roughly five times the cross-sectional area (πr^2) of a 2.1-mm-i.d. column.

8. Set gradient conditions on the chromatograph. A recommended gradient protocol for peptide mapping is as follows:

0-5 min	10% B
5-65 min	10%-60% B
65-70 min	60%-90% B
70-80 min	90% B.

BCU IASI CENTRAL UNIVERSITY LIBRARY

Gradient rate defines the change in the gradient over time. HPLC gradients are controlled by the pump dispensing buffer A, which can vary from 0% to 100% of the total flow. The amount of buffer A plus the amount of buffer B must equal 100%. Thus 0% B means that the B pump is off and only the A pump is dispensing buffer, i.e., 100% A; and 100% B means that only the B pump is running.

9. Set detector settings at 215 nm primary trace and 280 nm secondary trace.

   A single-channel HPLC apparatus is perfectly acceptable for this procedure; monitoring should be at 215 nm.

10. Begin the HPLC analysis. Perform a blank run with buffers alone, a second run with the enzyme-only (or chemical cleavage agent–only) control, and finally the run with the protein sample (with a separate run for the native-protein control, if available). Inject 100 µl of each sample to start run and do not start gradient until the sample has had time to flow onto the column. Equilibrate with HPLC buffer A between runs.

    A 100-µl sample will take 0.66 min to load at a flow rate of 150 µl/min.

    The injection of the sample initiates data acquisition.

11. Compare traces at all wavelengths.

    See Anticipated Results for a discussion of data analysis.

PEPTIDE MAPPING BY HIGH-RESOLUTION SDS-PAGE

ALTERNATE PROTOCOL

This method has less resolving power than HPLC, but allows a comparison of the recombinant product with very small amounts (in the 10 to 100 ng range) of purified native protein using sensitive stains (such as silver or gold staining) or radiolabeling as a detection method. In addition, gel electrophoresis equipment is more commonly available in laboratories.

Additional Materials (also see Basic Protocol 1)

2× SDS sample buffer (*UNIT 10.1*)
16 × 16–cm × 1-mm or 16 × 20–cm × 1-mm Tris-tricine slab gel (*UNIT 10.1*)
85°C water bath or heating block
XAR X-ray film or equivalent (optional; for detection of radiolabeled proteins, if used)
Additional reagents and equipment for radiolabeling of proteins (*UNITS 3.3*; optional), proteolytic digestion (*UNIT 11.1*) and/or chemical cleavage (*UNIT 11.4*) of proteins in solution, digestion in SDS-polyacrylamide gel (optional; see Support Protocol), one-dimensional SDS-PAGE (*UNIT 10.2*), and detection of proteins in gels (*UNIT 10.5*) or on membranes (*UNIT 10.8*)

Prepare protein sample

1. Prepare the sample as for peptide mapping by HPLC by digesting either in solution (see Basic Protocol 1, steps 1 to 5) or in SDS-polyacrylamide gel (see Support Protocol, steps 1 to 9).

   It is important to prepare the recombinant and native proteins in the same way. Always include an enzyme-only control.

   Start with 25 to 50 pmol protein if silver staining is to be used as the detection method. For Coomassie staining, start with 150 to 300 pmol protein. For more sensitive detection, proteins in the 1 to 5 pmol range can be visualized by gold staining (UNIT 10.8) or radiolabeled before digestion with ^{125}I by the iodogen method (Markwell and Fox, 1978; UNIT 3.3).

BOULAS/CENTRAL UNIVERSITY LIBRARY

2. Add 1 vol of 2× SDS sample buffer to each sample. Heat 5 min at 85°C and then load into wells formed in a high-resolution Tris-tricine slab gel.

For this method to work, it is important to use Tris-tricine gels, as regular Laemmli gels do not provide the required resolution of small fragments. If the sample was prepared by digestion in a gel slice (Support Protocol) and already contains some SDS, then simply calculate how much 2× SDS sample buffer needs to be added to provide a final concentration of 2% SDS. If the sample was precipitated with acetone, redissolve it in 10 µl of 2× SDS sample buffer and add 10 µl water.

3. Perform one-dimensional electrophoresis as indicated in *UNIT 10.1*.

Bromphenol blue accompanies fragments in the 3-kDa range. It is wise to stop the run when the bromophenol blue band is at least 1 cm from the bottom of the gel.

4. Prepare gel for staining as recommended in *UNIT 10.5* for the particular stain (silver or Coomassie blue) that is to be used.

With silver staining, best results are obtained if the gel is washed with at least four changes of 50% methanol over 2 hr or equilibrated with 50% methanol overnight. Small peptides do not stain as intensely as larger peptides and may leak out of the gel. Fixing with glutaraldehyde (UNIT 10.5) may prevent this from happening and will enhance the staining obtained with silver.

5a. Detect proteins by standard methods, such as silver or Coomassie blue staining (*UNIT 10.5*).

5b. If proteins have previously been radiolabeled, fix and dry the gel, then detect proteins directly using a phosphoimager *or* by autoradiography (for ^{32}P or ^{125}I) or fluorography (for ^{3}H, ^{14}C, or ^{35}S) with XAR X-ray film.

Methods for autoradiography and fluorography can be found in Laskey and Mills (1975).

SUPPORT PROTOCOL

DIGESTION OF PROTEIN IN AN SDS-POLYACRYLAMIDE GEL SLICE

On occasion, it is desirable to compare a recombinant protein with authentic material after purification by SDS-PAGE. For this protocol, ~50 pmol protein is required. The samples are subjected to SDS-PAGE in the normal manner (*UNITS 10.1 & 10.4*) and stained with Coomassie blue (*UNIT 10.5*). Bands of interest are simply cut from the gel and treated as described below; the digested fragments can then be analyzed by HPLC (Basic Protocol 1) or by high-resolution gel electrophoresis (Alternate Protocol).

***Additional Materials** (also see Basic Protocol 1)*

10 mM Tris·Cl (pH 7.5)/0.1% (w/v) SDS
1.5-ml Kontes microcentrifuge tube and pestle, or equivalent
1.5-ml low-protein-binding microfilterfuge tube (0.2-µm cellulose acetate; Rainin or Costar)
Dry ice/ethanol bath (optional)
Additional reagents and equipment for one- or two-dimensional SDS-PAGE (*UNITS 10.1 & 10.4*), staining with Coomassie blue (*UNIT 10.5*), and proteolytic digestion of proteins in gels (*UNIT 11.3*)

Gel purify the protein of interest

1. Load a sample containing ≥50 pmol protein onto an SDS-polyacrylamide gel and perform one- or two-dimensional SDS-PAGE (*UNITS 10.1 & 10.4*), then stain with Coomassie blue and destain in standard fashion (*UNIT 10.5*).

The following list of equivalences is helpful for determining the required amount of protein: for a 100-kDa protein, 100 µg = 1 nmol; for a 50-kDa protein, 50 µg = 1 nmol;

BRU IASI / CENTRAL UNIVERSITY LIBRARY

and for 25-kDa protein, 25 ug = 1 nmol. Because proteins are reduced and denatured in SDS sample buffer, it is not generally necessary to further treat the protein with reducing or alkylating agents. However, if there is concern about disulfides, the protein can be reduced and alkylated prior to electrophoresis.

2. Wash the gel slice in water, using four changes in 1 hr with constant shaking, to remove as much of the destaining solution as possible.

3. Dry the gel slice in air for 15 min.

Digest or chemically cleave the sample

4. Rehydrate the gel slice with digestion buffer containing an appropriate amount of the selected enzyme or chemical cleavage agent: e.g., 140 µl digestion buffer containing 0.003 U endoproteinase Lys-C per µg protein, or ~1:25 (w/w) enzyme to protein.

   Refer to UNITS 11.1 & 11.3 for protocols for different proteolytic enzymes, including appropriate digestion buffers. This methodology can also be used to cleave proteins using chemical agents such as CNBr, which cleaves at Met groups; BNPS-skatole, which cleaves at Trp groups; or hydroxylamine, which cleaves at Asn-Gly bonds. Procedures for using these agents can be found in UNIT 11.4.

5. Macerate the gel with a tight-fitting pestle (a Kontes microcentrifuge tube and pestle are ideal). Carefully rinse the pestle with an additional 50 µl digestion buffer.

6. Cap the tube, wrap a strip of Parafilm tightly around the cap to prevent evaporation, and allow to digest overnight (~16 hr) at 37°C.

7. After digestion, separate the protein fragments from the gel pieces by centrifuging 5 min at ~2000 to 4000 × *g* (see manufacturers' recommendations) through a 1.5-ml low-protein-binding microfilterfuge tube.

   An efficient way to transfer the gel mixture to a microfilterfuge tube is to simply cut the bottom off the microcentrifuge tube containing the digest with a microcentrifuge cutter (Fisher) and assemble onto a microfilterfuge tube. The whole unit (cut microcentrifuge tube and microfilterfuge tube) can be stabilized by assembling it inside a 10-ml round plastic disposable tube. The unit is then centrifuged 5 min at low to medium speed (e.g., ~2000 to 4000 × g; see manufacturers' recommendations) in a tabletop centrifuge to allow all gel pieces to flow into the top compartment of the microfilterfuge. The use of a micropipet to transfer the mashed gel pieces to the microfilterfuge is not recommended. The use of a micropipet is not recommended because gel pieces will adhere to the inside of the pipet tip, reducing fragment recovery.

   Extracting the protein fragments from the gel is critical for good resolution of fragments by high-resolution gel electrophoresis.

8. Reextract the gel pieces with 100 µl of 10 mM Tris·Cl (pH 7.5)/0.1% SDS for 20 min at room temperature and then recentrifuge through the microfilterfuge tube.

9. Combine the extracts (containing the peptide fragments).

   If samples are to be analyzed by high-resolution SDS-PAGE, these extracts can be applied directly to the gel and steps 10 to 12 of the present protocol should be omitted; instead, proceed directly to electrophoresis (see Alternate Protocol, step 2).

Acetone precipitate samples that are to be analyzed by HPLC

10. Add 9 vol of –20°C acidified acetone to sample. Allow to precipitate overnight at –20°C or 2 hr in a dry ice/ethanol bath.

11. Microcentrifuge sample 15 min at maximum speed, 4°C. Carefully decant supernatant from the pellet.

BOULDER / CENTRAL UNIVERSITY LIBRARY

12. Dry the pellet in a Speedvac to remove all traces of acetone. Resuspend pellet in HPLC buffer A and proceed to separation by HPLC (see Basic Protocol 1, step 7).

Acetone precipitation is necessary to remove excess SDS present when digestion is performed in a gel slice. SDS levels should be reduced below 0.05% before the sample is applied to the HPLC column.

BASIC PROTOCOL 2

DETERMINING THE DISULFIDE LINKAGE PATTERN OF A RECOMBINANT PROTEIN

Correct formation of the disulfide linkages of a recombinant protein is at least as important for protein function as correct amino acid sequence, but can sometimes be more difficult to establish. First, the sample is treated to alkylate any free thiol groups so as to block their ability to catalyze disulfide interchange ("scrambling"); this is done using denaturing but nonreducing conditions to expose any buried thiols. In the next, key step the protein is digested by chemical or enzymatic means in such a way as to cleave the polypeptide chain at least once between successive Cys residues (except where two consecutive Cys residues share a disulfide bond, when this is not necessary). Resulting fragments containing only a single disulfide bond are then isolated by reversed-phase high-performance liquid chromatography (RP-HPLC) and identified by a combination of mass spectrometry and Edman analysis; each bond will either be holding two small peptide chains together, or holding a single polypeptide chain in a loop. This apparently simple task is often frustrated by the presence of closely spaced Cys residues (because it is necessary to cleave between them) and by the difficulty in ensuring that disulfide scrambling does not occur.

Given the huge variety of possible protein properties, it is impossible to give a single recipe for determining disulfide linkage patterns that can be blindly followed in all cases. The protocol that follows takes protein idiosyncracies into account as much as possible in the annotations which follow (or sometimes precede) critical steps; the proteins granulocyte/macrophage colony-stimulating factor (GM-CSF) and macrophage colony-stimulating factor (M-CSF) are discussed as examples.

NOTE: All reagents used should be of analytical grade or better. Solutions should be made fresh except where otherwise stated (see Reagents and Solutions). Special precautions are necessary when working with iodoacetic acid, concentrated formic acid, cyanogen bromide, and trifluoroacetic acid (TFA): these precautions are described at the points where use of these chemicals is described.

Materials

Sample of recombinant protein
Alkylation buffer (see recipe)
Nitrogen or argon, oxygen-free
Guanidine hydrochloride
50 mM and 2 M ammonium bicarbonate
2 mg/ml cyanogen bromide in 70% formic acid (see recipe)
HPLC solvents A and B for disulfide linkage analysis (see recipe)
1% formic acid
Proteases: pepsin, *Staphylococcus aureus* V8 protease (endoproteinase Glu-C), and/or trypsin (porcine variety, Sigma)
Glacial acetic acid
1 M dithiothreitol (DTT)

Pretreated dialysis tubing with appropriate MWCO (see recipe)
Apparatus for reversed-phase high-pressure liquid chromatography (RP-HPLC), preferably with automatic sample injector

BGU IASI UCENTRALNY UNIVERSITY LIBRARY

Appropriate analytical HPLC column: e.g., 25 cm × 4 mm i.d. filled with Nucleosil 5 µm, 300 Å C8 particles (Macherey-Nagel)

Appropriate preparative HPLC column: e.g., 25 cm × 4 cm i.d. filled with Nucleosil 5 µm (as above) for purifying up to ~0.2 mg protein, or 25 cm × 10 mm i.d. with same packing for up to ~1.5 mg protein (see Critical Parameters for further guidance)

Vacuum centrifuge: e.g., Speedvac (Savant)

Equipment for mass spectrometry of peptides

Additional reagents and equipment for dialysis (*UNIT 4.4*; *APPENDIX 3B*)

Block free thiols

1. Dissolve the protein to ~0.1 mM in alkylation buffer and incubate 30 min at room temperature (~22°C) in the dark under nitrogen or argon.

   Use 0.2 to 2 mg protein (depending on availability) for steps 1 to 3 in the first instance, and repeat on a larger scale if necessary.

   This step serves to alkylate any easily accessible thiol groups. Radiolabeled alkylating reagent (e.g., [^{14}C]iodoacetic acid) may be used to facilitate the later location of modified peptide.

2. Add solid guanidine hydrochloride to a final concentration of 6 M and continue incubation for a further 60 min.

   At this stage any previously buried thiols should become alkylated. Do not add thiol at the end of the incubation to consume excess alkylating agent in the normal fashion, as disulfides would then be at risk.

   It may be necessary to add the guanidine earlier on in order to solubilize the protein, in which case the 30-min preincubation should be omitted.

3. Transfer the sample to a prepared dialysis bag. Dialyze the sample (*UNIT 4.4; APPENDIX 3B*) in the dark at 4°C against 50 mM ammonium bicarbonate, using a bath volume ≥2000 times greater than the sample volume and stirring the bath magnetically. Change the solution after 1 hr, 2 hr, and 4 hr, and continue dialysis overnight (~18 hr total). Recover the protein by lyophilization.

   IMPORTANT NOTE: *Take great care not to damage the dialysis bag. A plastic automatic pipet tip is preferred for the transfer, but if a glass Pasteur pipet is used, remove sharp edges by heating it in a flame prior to taking up sample; avoid damage due to sharp fingernails, especially when knotting the bag.*

   Proteins known not to precipitate on removal of chaotrope can be recovered more quickly by gel filtration than by dialysis. Furthermore, steps 1 to 3 are usually omitted when it is known from attempted reaction with thiol-specific reagents or radiolabeled reagents that there are no unpaired Cys residues. Whenever there is reason to suspect the presence of an unpaired Cys—e.g., an odd number of Cys residues in the sequence, reaction with thiol-specific reagents (Ellman, 1959; Thannhauser et al., 1984) or radiolabeled reagents —these blocking steps must precede the fragmentation procedures. For example, blocking was necessary for the protein G-CSF (Wingfield et al., 1988) but not for GM-CSF (Schrimsher et al., 1987).

Fragment the polypeptide chains

4. Depending on the known or expected amino acid sequence, select digestion conditions (cyanogen bromide, one or more proteolytic enzymes, or a combination of these in succession) that will cleave the polypeptide chain at least once between successive Cys residues (see Critical Parameters for guidance). Perform an appropriate combination of the treatments described in steps 5 to 8.

   If the protein is known to contain a few well-spaced Met residues, cyanogen bromide is often useful as a first fragmentation step.

B. C. U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

For cyanogen bromide treatment

5. Dissolve the recombinant protein at 1 mg/ml in 2 mg/ml cyanogen bromide/70% formic acid. Seal the tube and incubate 24 hr in the dark at room temperature.

   CAUTION: *Cyanogen bromide is toxic and volatile; concentrated formic acid is very corrosive (see recipes for detailed precautions).*

6a. If the complexity of the analysis is high (i.e., a large protein and/or many disulfide bridges are involved) and the expected sequence suggests the problem can be simplified by working individually with the cyanogen bromide fragments, dilute the cyanogen bromide digest at least three-fold with HPLC solvent A and separate by RP-HPLC (steps 9 and 10), then proceed to subdigest the large fragments enzymatically.

6b. Otherwise, simply remove solvent and reagent under high vacuum in a vacuum centrifuge (avoid exposing other samples in the vacuum centrifuge to reagent vapors), then proceed to enzymatic digestion.

   Here and throughout the protocol, use the vacuum centrifuge with the heater switched off to lessen the risk of such reactions as deamidation, oxidation, and N-terminal cyclization.

For protease treatment

7a. *Using pepsin:* Dissolve the protein or protein fragments at ~1 mg/ml in 1% (v/v) formic acid, add pepsin to obtain a 1/100 (w/w) enzyme/substrate ratio, and incubate 16 to 24 hr at room temperature.

   Pepsin was useful for G-CSF, although trypsin or V8 protease were more useful for GM-CSF (Schrimsher et al., 1987).

7b. *Using trypsin or V8 protease:* Dissolve the protein or protein fragment at ~1 mg/ml in 0.1 M phosphate buffer, pH 7.0. Add enzyme to obtain a 1/100 (w/w) enzyme/substrate ratio, and incubate 4 hr at 37°C.

   When using trypsin, the porcine variety (Sigma) is preferred over the bovine enzyme as it retains its activity at lower pH.

8. If trypsin or V8 protease is used for digestion at neutral or slightly alkaline pH, acidify the digest to pH 3 with acetic acid just prior to HPLC (step 9) to stop the digestion and prepare the sample to be placed in the automatic injector.

   A peptic digest is already acidic and is injected directly; if it sits in the automatic injector, the digestion will continue.

Analyze digest by RP-HPLC

9. Subject the cyanogen bromide or enzyme digest to analytical RP-HPLC—e.g., using the trifluoroacetic acid/acetonitrile RP-HPLC system (*UNIT 11.6*), which is very convenient because both solvents are completely volatile. Use a gradient of 2% HPLC solvent B/min up to 100% B. Monitor at low wavelength (e.g., 214 nm), as many fractions may not contain aromatic residues.

Perform preparative HPLC

10. Perform preparative HPLC separations on digests that showed a suitable number of well-separated fragments in step 9. For best results, use a much shallower gradient (e.g., 0.2% solvent B/min) over the percentages of interest deduced from the screening run.

    For further discussion of the design of the preparative gradient, see, for example, Rose et al. (1992).

BSU IASI CENTRAL UNIVERSITY LIBRARY

Particularly for preparative HPLC runs, it is preferable to acidify the digest (if not already acid) to pH 3 with acetic acid prior to injection (check by spotting onto pH paper), as otherwise the early part of the chromatogram may be compromised. Ideally, choose column dimensions and flow rate compatible with the quantity of digested protein (see Critical Parameters). If several milligram of protein are available, it makes life much easier to work at the 1.5-mg scale, but remember that it may sometimes be necessary to return to step 7 and continue the analysis with a different digest.

11. Recover peptides by drying the collected fractions in the vacuum centrifuge, then redissolve each in 200 µl HPLC solvent A. Divide each fraction, placing two analytical-sized aliquots (10 µl each if a major peak and >0.5 mg digest were used; 30 µl if a minor peak and <0.5 mg digest were used) directly into the automatic injector sample tubes, and retaining the remainder separately.

The original chromatogram obtained during collection of the corresponding preparative fraction should be used to select an appropriate sample volume and attenuation level for the analytical runs that will follow.

Locate fractions containing disulfide bonds

12. To one sample add 2 M ammonium bicarbonate to 50 mM final and 1 M DTT to 50 mM final. To the second sample (control) add ammonium bicarbonate but no DTT. Close tubes and incubate 2 hr at 37°C.

Steps 12 and 13 constitute the simplest of several possible ways of locating fragments containing a disulfide bond: attempted reduction of disulfide bonds in each HPLC fraction with DTT followed by reinjection on a shallow gradient.

To perform this assay on a dry sample, the peptide can simply be dissolved in 50 mM ammonium bicarbonate/50 mM DTT.

13. Acidify the samples with a few microliters of acetic acid (test a blank with pH paper), agitate to liberate the carbon dioxide bubbles, then analyze the samples in pairs (reduced and nonreduced) by RP-HPLC, using the analytical column but with the shallow gradient used for the preparative run (step 10).

As mentioned in step 11, the original chromatogram obtained during collection of the corresponding preparative fraction, and the proportion taken for analysis, permit an appropriate choice of attenuation to be made for these analytical runs.

14. Interpret the pairs of chromatograms (attempted reduction and control). When all has gone well, if no disulfide bond is present there should be essentially no difference between attempted reduction and control. If a disulfide bond is present, then either this bond is holding two small peptide chains together, or it is holding a single polypeptide chain in a loop. In the former case, two peaks should be visible on reduction; in the latter case all that will be observed is a small shift to later retention time (opening of the loop exposes more of the peptide chain to the stationary phase).

The reason for performing a control incubation in the absence of reductant is that some fractions may shift retention time or develop structure due to changes such as opening of a homoserine lactone, cyclization of an N-terminal Gln, or oxidation of methionine to its sulfoxide, and this change might otherwise be mistaken for reduction.

Other sources of difficulty in interpreting results include the following: One of the components liberated by reduction of a disulfide bond may elute early and be difficult to detect because of the presence of salt and DTT peaks. One component may be short and devoid of aromatic residues and so be much less visible on the chromatogram. Finally, because only one HPLC dimension was used for the initial collection, there may be a nonreducible component present in addition to a reducible one; this need not be a problem when both Edman and mass spectrometric techniques are used for characterization.

BOULDER / CENTRAL UNIVERSITY LIBRARY

Characterize Cys-containing peptides

15. For each fraction identified as containing a disulfide bond, repeat the reduction (step 12) with an appropriate portion of the remaining nonreduced fraction (set aside in step 11), repeat the HPLC separation, and collect the reduced components.

    For an example of where this repeated reduction and separation strategy was used, see Schrimsher et al. (1987). To determine sample size, the mass spectroscopist who will perform that analysis should be consulted regarding how much protein will be required.

16. Divide the fractions and submit these reduced components, together with some of the unreduced fraction, for examination both by mass spectrometry and by Edman degradation.

    The combination of mass spectrometry and Edman degradation provides the quickest and surest approach to characterizing the fractions containing disulfide bonds. Mass spectrometry provides the molecular weight of the peptide component (or components) present, while Edman analysis should give their N-terminal sequences.

    The expense and sophistication of mass spectrometric equipment is such that usually, where it exists, a competent person is present to give advice on its use. Electrospray ionization mass spectrometry (ESI-MS) is most suited to this application, although good results may also be obtained by fast-atom-bombardment (liquid secondary-ion) mass spectrometry if a high-performance machine is available. Matrix-assisted laser-desorption ionization time-of-flight (MALDI-TOF) mass spectrometry may also be used (see UNIT 11.6); mass precision is generally lower than with fast-atom-bombardment or ESI spectrometry, but can be adequate if the time-of-flight analyzer is operated in reflectron mode and carefully calibrated (using an internal standard), and is excellent if the MALDI source is connected to a Fourier-transform ion cyclotron resonance analyzer.

    Some peptide components may not be susceptible to Edman degradation. A peptide may be N-terminally blocked (e.g., by acetyl for the N-terminal peptide of the protein, or by a residue of pyroglutamic acid if a Gln residue liberated during proteolysis has cyclized). This problem is discussed in Critical Parameters and Troubleshooting.

17. Align the peptides identified in step 16 with the known sequence of the recombinant protein. Interpret the molecular weights determined quite precisely by mass spectrometry.

    Even if the Edman degradation did not provide sequence information up to the C terminus of a given peptide, the N-terminal starting point should be unambiguous: the C terminus can then be deduced from the mass of the peptide. Taken together, the Edman and mass spectrometry data generally allow the disulfide bonding pattern of the component peptides to be deduced, and hence the bonding pattern of the recombinant protein. Even where this is not the case, sufficient information is gained to greatly facilitate a second attempt.

    In some cases, the fragments identified indicate neither a single disulfide bond in a loop nor two short chains held together by one disulfide bond. In such cases (e.g., where two chains are both linked to a third by disulfide bonds; Rose et al., 1992), further cycles of fragmentation, isolation, and identification must be performed in order to cleave between the Cys residues.

    Scintillation counting of fractions permits rapid location of any Cys residues radioalkylated in Step 1. Characterization of any such fractions by Edman degradation and mass spectrometry completes linkage assignment by proving that the Cys residues so alkylated could not have been involved in disulfide bonds.

BRU IASI / BETRAUNIVERSITY LIBRARY

DETERMINING CHARGE HETEROGENEITY BY TWO-DIMENSIONAL TITRATION CURVE ANALYSIS

Unfolding of proteins can be monitored by running the sample across a urea gradient—the resulting denaturation curve, often sigmoidal in shape, represents the transition from the folded to the unfolded state. In general, pH/mobility curves of proteins are generated in a polyacrylamide IEF slab gel (see Fig. 7.3.1 for an outline of the procedure).

Materials

Acrylamide/bisacrylamide stock solution (see recipe)
Carrier ampholytes: e.g., Pharmalyte (Pharmacia Biotech), Bio-Lyte (Bio-Rad), or Servalyte (Serva)
10% (w/v) ammonium persulfate stock solution (see recipe)
N,N,N′,N′-tetramethylethylene diamine (TEMED)
Kerosene
Electrode solutions (see recipe)
Sample of recombinant protein in aqueous solution containing ≤10 mM Tris·acetate (or other salt or buffer consisting of weak acids and bases)
Coomassie blue staining solution: 0.1% (w/v) Coomassie brilliant blue R-250/25% (v/v) ethanol/8% (v/v) acetic acid
Destaining solution: 25% (v/v) ethanol/8% (v/v) acetic acid

Gel Bond PAG sheets
Electrophoresis unit for horizontal slab gel operation: e.g., Multiphor (Pharmacia Biotech)
Levelling table
IEF electrode filter paper strips (Pharmacia Biotech)

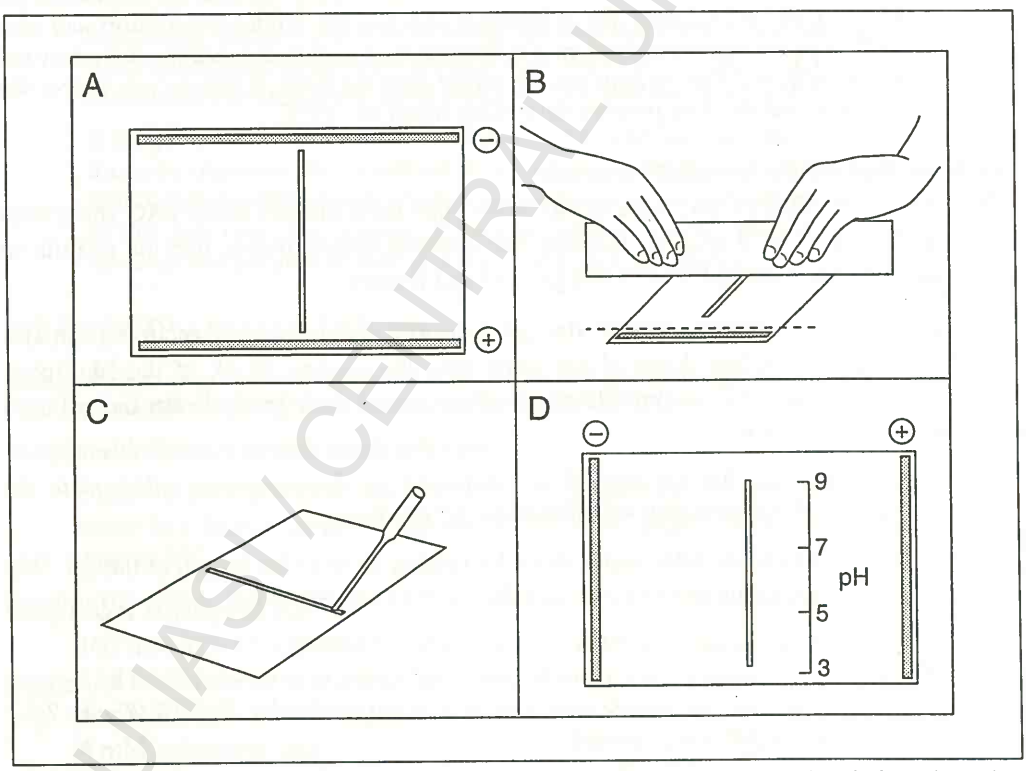


Figure 7.3.1 Experimental procedure for generating titration curves by isoelectric focusing. First the pH gradient is formed by focusing the carrier ampholyte mixture without sample (first-dimension run; **A**); the electrode strips and the gel layer underneath are then removed (**B**); the sample is applied in the trench now perpendicular to the pH gradient (**C**); the second-dimension run is performed perpendicular to the first dimension axis (**D**). From Righetti et al. (1979).

FUJIASI / CENTRAL UNIVERSITY LIBRARY

Constant-power supply (up to 3000 V, 200 mA, 20 W)
Thermostatic water bath: e.g., MultiTemp (Pharmacia Biotech)
Titration curve kit (Pharmacia Biotech)

Additional reagents and equipment for denaturing IEF in slab gels (*UNIT 10.2*)

Cast the gel

1. Cut a Gel Bond PAG sheet in half (to 12.5 × 12.5 cm) to form a gel-supporting foil. Roll the supporting foil on the glass plate of the electrophoresis unit, putting a few drops of water on the back to secure adhesion.

   The hydrophilic surface of the foil should face the operator; this is the surface designed for adhesion of the gelling layer.

2. Complete the cassette by placing the glass plate–foil sandwich between the two marks on top of the slot former and fastening them together using four clamps. Lay out the completed cassette horizontally, preferably on a levelling table.

3. Prepare 28 ml of an appropriate gel solution—e.g., containing 5%T and 4%C monomers and 2% ampholytes—from acrylamide/bisacrylamide stock solution and carrier ampholytes (see *UNIT 10.2* for details of gel casting). Degas for 5 to 10 min.

   Preparing 28 ml of solution provides a slight excess, as 25 ml will be needed to fill up the gel cassette.

4. Add 4 µl ammonium persulfate stock solution and 1 µl pure TEMED per milliliter of gelling solution; mix carefully. Immediately inject the gel solution between the slot-former and the upper glass plate (the solution will fill the cassette by capillary action). Let stand while polymerization occurs (this will be complete after 1 hr).

   IMPORTANT NOTE: *When titration curves are run under denaturing conditions (in 8 M urea), the gel solution will polymerize very quickly. High temperatures will also greatly accelerate polymerization. If polymerization occurs too quickly, there may not even be sufficient time to pour out the liquid from the flask. If this occurs, repeat the procedure and this time precool the gelling liquid to ~10°C.*

Perform first-dimension electrophoresis

5. Remove the clamps and separate the glass plate from the Gel Bond PAG sheet with a gentle twist of a spatula. Holding the cassette upside down, use the spatula to introduce air bubbles between the gel and slot former.

6. Starting from a corner, remove the gel from the plastic cover containing the slot former. Spread a few drops of kerosene onto the cooling block of the Multiphor chamber (set at 10°C) to provide good adhesion, and then gently lower the gel onto the cooling surface.

   Any trapped air bubbles should be eliminated, as their presence will lead to the formation of hot spots that will distort the pH gradient.

7. Prepare IEF electrode filter paper strips by cutting them to the length of the gel. Wet in electrode solutions and blot excess solution with a paper tissue. Apply strips to gel surface.

   For good performance, the electrode strips and platinum wires should all be parallel to one another and the sample trench should be perpendicular. Refer to Figure 7.3.1 for this and the following operation.

8. Perform first-dimension electrophoresis at 10 W constant power with voltage and current set to a maximum (typically, at steady state, the voltage should be ~1000 V).

   This focusing step is usually terminated in 50 min.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Perform second-dimension electrophoresis

9. Remove the electrode strips and use a long blade to cut away the strips of gel underlying the electrode strips (see Fig. 7.3.1B).

   This step is important, as otherwise considerable heating will occur in these regions due to high conductivity.

10. Turn the gel around 90° on the cooling plate. Prepare and apply new electrode strips, using the same electrode solutions as for the first dimension. The sample trench is now parallel to the electrode strips (see Fig. 7.3.1).

11. Apply the sample by filling the trench with ~100 µl of protein solution (typically 100 to 400 µg protein, depending on sample heterogeneity, at 1 to 4 mg/ml; see Fig. 7.3.1C).

12. Perform second-dimension electrophoresis at 600 V for 5 to 20 min (based on sample mobility).

    High levels of salts, as well as salts and buffers comprising strong acids and bases, must be strictly avoided in any focusing process. For best results, the sample should not contain more than 10 mM salt or buffer, and these should consist of weak acids and bases—e.g., Tris-acetate, but not Tris·Cl or sodium acetate.

Stain and destain gel

13. Place the gel in Coomassie blue staining solution and allow to stain 1 hr with shaking.

    Staining solution can be reused at least 10 times. Store at room temperature protected from light.

14. Replace the staining solution with destaining solution. Destain with shaking, changing solution several times until a clear background is seen.

    Destaining solution can also be stored at room temperature and reused several times if filtered through a charcoal cartridge after each use. Beware of excessive evaporation of both ethanol and acetic acid, which will weaken the destaining process.

    If the gel is over-destained, the titration curve zones will also fade away. In this case, it can be restained and destained. When too little sample is available to be visualized in this fashion, any variant of silver staining may be used (Rabilloud, 1990; UNIT 10.5). In this last case, the amount of sample can be 20 times smaller. For best results, however, the gel slab should be thinner (e.g., 0.5 mm thick).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Acrylamide/bisacrylamide stock solution

Dissolve 28.7 g acrylamide and 1.3 g *N,N'*-methylene bisacrylamide in distilled water to a final volume of 100 ml (30%T, 4%C final). Store up to 3 weeks in the dark at 4°C.

Alkylation buffer, pH 8.0

100 mM Tris·Cl adjusted to pH 8.0 with 6 M HCl or glacial acetic acid
125 mM NaCl
5 mM EDTA
4 mM iodoacetic acid
Prepare fresh just before use

Concentrated stock solutions of EDTA (sodium salt; APPENDIX 2E) and iodoacetic acid in water may be made up and kept at –20°C (stable at least 1 month). Iodoacetic acid

continued

BOSTON UNIVERSITY LIBRARY

should be bought in small amounts and should take the form of white crystals. The solid and its solutions are light-sensitive and should therefore be shielded with aluminum foil and kept protected from direct sunlight.

CAUTION: *Solid iodoacetic acid and its concentrated solutions are corrosive—wear gloves when handling these materials.*

Ammonium persulfate, 10% (w/v)

Dissolve 50 mg ammonium persulfate in 0.5 ml distilled water. Prepare fresh each time.

Cyanogen bromide, 2 mg/ml in 70% formic acid

Weigh an empty open small glass tube or vial together with its cap, and then move this container to a chemical fume hood. Add a few small crystals of cyanogen bromide, replace the cap, and then reweigh the container to calculate the amount of cyanogen bromide added. Dilute with sufficient 70% formic acid to obtain a 2 mg/ml solution. (This procedure avoids the need to use a balance inside the chemical fume hood.) Prepare fresh.

CAUTION: *Cyanogen bromide, though a solid at room temperature, is quite volatile and sufficiently toxic to have been used as a chemical warfare agent (albeit a rather ineffective one). It should be manipulated in a chemical fume hood and gloves should be worn when handling.*

Buy good-quality reagent, keep solid protected from light at 4°C to minimize degradation, and make up fresh solution as required, as otherwise contaminating free bromine will tend to oxidize Met and Trp residues, which can result in cleavage at Trp and failure to cleave at Met.

Dialysis tubing, pretreated

Obtain dialysis tubing with appropriate MWCO (molecular weight cut-off) for the protein under study—e.g., Visking 18/32, which has a cut-off after boiling of <5000—and prepare an appropriate length of tubing fresh before use by treating it as follows to remove glycerol and residues from manufacture. Bring 1% sodium bicarbonate solution to a simmer in a beaker which has no chips or cracks (as these could pierce the tubing). Place the tubing in the solution and simmer 10 min; do not boil vigorously or the tubing will protrude from the solution and become brittle. Cool by adding distilled water, then thoroughly rinse the tubing inside and out with distilled water.

Excess prepared tubing can be stored in water or buffer for a few days at 4°C if precautions are taken to avoid microbial growth (e.g., add sodium azide to 0.005%).

IMPORTANT NOTE: *Throughout tubing preparation, beware of sharp edges, including fingernails and especially pipets, which can easily damage the tube and lead to complete loss of sample during the experiment.*

Digestion buffer

10 mM Tris·Cl, pH 8.0
0.1 mM EDTA
0.02 M NH_4HCO_3
0.1% SDS (include *only* if protein is to be digested in a gel slice)
Prepare fresh

Different proteolytic enzymes and chemical cleavage agents require different digestion buffers; refer to UNITS 11.1 & 11.4 for a discussion of these requirements.

DSU IASL CENTRAL UNIVERSITY LIBRARY

Electrode solutions

Use 50 mM solutions of either weak acids (e.g., acetic acid) or weak bases (e.g., Tris) as anolyte and catholyte, respectively (strong acids and bases should be avoided). Alternatively, use 50 mM solutions of isoionic zwitterions (e.g., Asp, pI 2.77, or Glu, pI 3.22) as anolyte and 50 mM Lys (pI 9.74) or 50 mM Arg (pI 10.76) as catholyte. Store in cold room to avoid bacterial growth. Keep basic solutions tightly capped and check the pH before use. If the pH decreases ≥0.5 pH units (due to CO_2 absorption), discard the solution.

HPLC solvents for disulfide linkage analysis

RP-HPLC solvent A: Add 1 g HPLC-grade trifluoroacetic acid (TFA; Pierce or ABI) to 1 liter of HPLC-grade water (e.g., purified using a Milli-Q system, Millipore) and mix. Degas either by vacuum filtration in an all-glass device (Millipore) or by sparging with helium. The solution is good for at least 3 months when stored at room temperature, protected from dust, and degassed prior to each use.

HPLC solvent B: Add 1 g TFA to 100 ml HPLC-grade water and mix. Dilute to 1 liter with HPLC-gradient-grade acetonitrile. Degas as for solvent A except, if using vacuum filtration, employ a solvent-resistant filter membrane (e.g., Millipore). The solution is good for at least 3 months when stored at room temperature, protected from dust, and degassed prior to each use.

CAUTION: *Undiluted TFA is very corrosive: wear eye protection (preferably a face shield) and gloves. Once diluted to 1 g/liter, it presents no danger.*

HPLC solvents for peptide mapping

HPLC solvent A: 0.1% (v/v) TFA in water
HPLC solvent B: 0.07% (v/v) TFA/80% acetonitrile

Filter and degas both solvents. For best results, balance absorption of solvents A and B at 215 nm by adding TFA until absorption values are equal. Store up to 1 week at room temperature.

COMMENTARY

Background Information

Peptide mapping

The number of different digests required to characterize a recombinant protein depends on the analytical method used and on the quality of information produced. Analysis by high-performance liquid chromatography (HPLC; Basic Protocol 1) allows the use of proteolytic agents that cut proteins generally into small fragments (such as trypsin, which cuts at Arg and Lys residues; endoproteinase Lys-C, specific for Lys residues; or endoproteinase Asp-N, specific for Asp residues). Analysis by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Alternate Protocol), on the other hand, requires larger fragments (≥2 kDa in size) and thus is limited to agents that cut the protein of interest at amino acid residues that are less frequent (such as CNBr or BNPS-skatole, which cut at Met and Trp residues, respectively) or to partial digests with proteolytic enzymes (such as endoproteinases Lys-C or Asp-N, which cut at Lys and Asp residues, respectively). These proteolytic agents have very strict specificities and easily predicted patterns of digestion. Other proteases have rather broad specificities (e.g., chymotrypsin is specific for Tyr, Phe, and Trp and to a lesser extent for Leu, Met, Ala, Asp, and Glu) and therefore produce a large number of fragments. These proteases are ideal for comparing a recombinant protein to the native protein because many more sites can be probed with only one digestion. Still other proteases, such as endoproteinase Glu-C (otherwise known as *Staphylococcus aureus* V8 protease), recognize one or two sites (in this case Glu and Asp) depending on the buffer conditions used for the digestion. For a useful discussion of these enzymes, see Allen (1989) and *UNIT 11.1*.

For recombinant proteins, the investigator has the advantage of knowing the predicted amino acid sequence and can design a specific strategy based on the distribution of key amino acids that define chemical and protease cleav-

BODLEIAN CENTRAL UNIVERSITY LIBRARY

age sites. Several computer programs predict cut sites, HPLC elution patterns, and molecular weights of fragments. The Peptide Sort program from Genetics Computer Group (GCG; University of Wisconsin, Madison, Wis.) is one such program. Because many cleavage agents give rise to a combination of complete and partial digests, it is important to treat the recombinant protein and the native isolated protein (if available) under the same conditions.

If it is impossible to obtain enough native material for comparison by the methods described above, there are a number of analytical techniques that can be used to prove the recombinant protein has the correct structure. The easiest and most cost-effective method is to analyze individual peaks from an HPLC run for their amino acid compositions; this can be compared to the predicted composition for each fragment. This method is most powerful when combined with N-terminal amino acid sequencing, which can identify each fragment collected by HPLC. The intact protein can also be submitted for N-terminal sequencing. With new instruments, it should be possible to identify as many as 50 to 100 residues directly from the N-terminus. On several occasions it has been possible to sequence up to 50% of a protein from a single digest (N. Denslow, unpub. observ.).

Newer analytical methods rely on mass spectrometry to analyze the digest of recombinant protein (Langen et al., 1993; Patterson, 1994). The mass accuracy of mass spectrometers is sufficient to identify a fragment directly. The preferred technique relies on the use of electrospray ionization (ESI)–quadrupole mass spectrometers, especially if the mass spectrometer is in-line with the HPLC. ESI-MS has a mass accuracy of 0.01%. Accurate data, however, can also be obtained from a matrix-assisted laser-desorption ionization time-of-flight (MALDI-TOF) mass spectrometer, which typically has a mass accuracy of 0.03%. In this case, HPLC fractions must be collected and mixed with a matrix for analysis.

Disulfide linkage analysis

Correct formation of the disulfide linkages of a recombinant protein is at least as important for function as correct amino acid sequence, but (as can be appreciated from the complexity of Basic Protocol 2) it is sometimes more difficult to establish. The key to success generally involves digestion of the protein by chemical or enzymatic means in such a way as to cleave the polypeptide chain at least once between successive Cys residues—except where two consecutive Cys residues share a disulfide bond, when this is not necessary. This apparently simple task is often frustrated by the presence of closely spaced Cys residues, because it is not always easy to cleave between these. Some instructive examples showing the difficulties encountered when two Cys residues are very close or juxtaposed are discussed by Morris and Pucci (1985); Hojrup and Magnusson (1987); and Rose et al. (1992). Cleavage techniques are somewhat limited by the need to prevent disulfide scrambling (Hojrup and Magnusson, 1987; Rose et al., 1992). Scrambling can occur in the presence of strong mineral acid, and also in neutral to basic medium, which is why mildly acidic conditions are preferred where possible. Scrambling is catalyzed by the presence of free thiol; for this reason any unpaired Cys residues (which may be present in a small fraction of improperly refolded material) must be blocked in the first step.

It must be emphasized that people working with a recombinant protein usually have two enormous advantages over those working with a rare polypeptide isolated from natural sources—they have much more material, and they know most, if not all, of the amino acid sequence. These advantages must be exploited to the full: it is advisable to use trial amounts, change tactics if solubility problems are insurmountable, and design the fragmentation strategy around the known sequence. The method described is designed to give results as reliably and rapidly as possible with recombinant proteins, and would need extensive modification to be suitable for dealing with trace amounts of proteins derived from natural sources.

Titration curve analysis

Protein titration curves (in reality, pH/mobility curves) became suddenly popular in the late seventies after an original report by Rosengren et al. (1977), who demonstrated a simple method to force a protein to migrate across a quasi-stationary pH gradient in such a way as to trace a sigmoidal path resembling a classical potentiometric titration curve. These pH/mobility curves were found to be quite useful in a number of applications: (1) for screening of genetic mutants (Righetti et al., 1978a; Righetti et al., 1979); (2) for studying macromolecule-ligand interactions (Krishnamoorthy et al., 1978; Constans et al., 1980); (3) for demonstrating macromolecule-macromolecule interactions (Righetti et al., 1978b; Lostanlen et al., 1980; Gianazza et al., 1980); (4) for direct pK

determination of simple cations and anions and of uni-uni-valent amphoteric molecules (Righetti et al., 1979; Valentini et al., 1980); (5) for determining dissociation constants (K_d) of ligands and proteins and their pH dependence (Ek et al., 1980; Ek and Righetti, 1980); (6) for detecting protein and peptide microheterogeneity (Arnaud et al., 1980; Henner and Sitrin, 1984; Picard et al., 1987; Van Den Oetelaar and Hoenders, 1987); and (7) for choosing optimal pH conditions in ion-exchange chromatography (Fägerstam et al., 1983).

The times must have been ripe for such an idea, because at just about the same time, Creighton (1979) reported a method for studying the unfolding of proteins by running them across a urea gradient: these denaturation curves, often sigmoidal in shape, represent the transition from the folded to the unfolded state. In an analogous manner, it is possible to study the melting profiles of DNA (from double helix to single strand) by running them across a thermal gradient: these curves too are sigmoidal in shape (Riesner et al., 1991).

In general, pH/mobility curves of proteins are generated in a polyacrylamide slab gel. However, a number of modifications have been reported, including the use of: (1) cellulose acetate as support (Vickers and Robinson, 1983); (2) agarose thin layers (Troungos et al., 1982); (3) highly cross-linked (highly porous) polyacrylamide matrices (Bianchi-Bosisio et al., 1980); (4) thin-layer silica gels (Stahl and Muller, 1981)—in this last instance, however, no electric field is used and the analysis performed is a standard ascending thin-layer chromatography against a pH gradient created by buffer diffusion; (5) paper sheets as supports (Tate, 1981)—here too the pH gradient is obtained by bathing the paper sheet in a number of buffers titrated at different pH values; and (6) paper sheets in water-alcohol solvents (Jokl et al., 1979)—again, the pH gradient is obtained as in Tate (1981).

The typical gel slab size normally employed consists of a polyacrylamide plaque 12.5 × 12.5 cm and 1.6 mm in thickness. Because this is the typical plate commercially available, that is the only method described extensively in Basic Protocol 3. However, it must be emphasized that by using the Pharmacia Biotech Phast-System, titration curves can be run in a small format (typically 4 cm × 4 cm × 0.5 mm) and the entire operation (including staining and destaining) can be performed in only 50 min. Additionally, Rüchel and Trost (1981) have described a minichamber only 35 × 15 × 0.5 mm vertically inserted into two cylindrical buffer reservoirs. The gel slab is polymerized by using a two-vessel gradient mixer with two limiting solutions titrated to different pH values; thus, in this case too an artificial pH gradient (utilizing diffusion of nonamphoteric buffers) is created instead of a natural one (created by a current with the aid of a multitude of amphoteric buffers).

Critical Parameters and Troubleshooting

Peptide mapping

Excess SDS (>0.01%) is a problem for peptide mapping by HPLC (Basic Protocol 1). For samples digested in a gel slice in the presence of 0.1% SDS, the detergent must be removed prior to applying the sample to the HPLC. Otherwise, it will interfere with the resolution of fragments and the retention time on the column, and numerous artifactual peaks will be observed. SDS can be removed to a certain extent by precipitation in acidified acetone. Total removal of SDS from a protein solution requires either binding to an ion exchanger or electrophoresis. Ion-exchange precolumns that can be put in-line with the reverse phase column to remove SDS from a contaminated sample are available commercially (e.g., from The Nest Group).

Excess of either SDS or salt causes problems for high-resolution SDS-polyacrylamide gels. These factors should not be a problem if the method described in the Alternate Protocol is used. In the event that a protocol containing high salt or excess SDS is used to cleave a protein, levels of both can be reduced by precipitation with acidified acetone. Generally, proteins that are precipitated by acidified acetone in the presence of SDS are readily solubilized after precipitation, probably because a small amount of SDS adheres to the protein.

It is important to use the correct digestion buffer for each protease (Allen, 1989). If the wrong conditions are used, the enzyme is likely to lose its specificity, if it works at all. Some proteases can withstand up to 0.1% SDS or 10% acetonitrile without loosing activity. Others are sensitive to 0.01% SDS, and in these cases the conditions described here may not be suitable. Refer to *UNITS 11.1 & 11.4* for specific protocols. For best results, proteases should be of sequencing grade.

For high-resolution SDS-PAGE of proteolytic digests done in gel slices, it is crucial to

ROSWELL PARK CANCER INSTITUTE LIBRARY

separate the polypeptide fragments from the gel pieces prior to running the sample. As indicated in the protocol, this can be done easily by centrifuging the sample through a microfilterfuge. Recoveries are in the 80% range for this step. Failure to remove the gel pieces will widen each peptide band, giving rise to poor resolution in the gel.

Gels that will be silver stained should be washed with 50% methanol (four changes in 2 hr, or equilibration overnight) to remove substances that interfere with the staining process.

Disulfide linkage analysis

Most steps are critical when working with a given protein for the first time, as one can lose it by precipitation or adsorption to columns, or in various other ways! Usually, however, many properties of a protein (e.g., solubility and stability) are discovered during its isolation and purification: it is important to find out all that is already known about the protein before attempting to determine the disulfide bridge positions. Unless the protein is available in great abundance, it is advisable not to commit >0.1 to 0.2 mg to any untried step.

Blocking steps 1 to 3 must be performed whenever there is reason to suspect the presence of an unpaired Cys (e.g., an odd number of Cys residues in the sequence or reaction with thiol-specific reagents: Ellman, 1959; Thannhauser et al., 1984). These blocking steps must precede the fragmentation procedures to prevent scrambling (see Background Information).

Good HPLC technique is critical to success: the appearance of spurious peaks on chromatograms is often due to poor technique. The analytical column should be run at a flow rate of 0.6 ml/min. This column may also be used for preparative purposes with digests of up to ~0.2 mg protein: use flow rates of 0.6 ml/min when working with up to 0.1 mg digest and 1 ml/min when working with 0.1 to 0.2 mg digest. If working with larger amounts (up to ~1.5 mg digested protein), use a 25-cm × 10-mm-i.d. column filled with the same packing and operate at 3 ml/min (for 0.5 to 1 mg digest) or 4 ml/min (for 1 to 1.5 mg digest). When no longer required, columns must first be flushed with five column volumes of pure acetonitrile; flow may then be stopped. If the flow is stopped with HPLC solvent A or B present, the column will rapidly deteriorate and silica or stationary phase may block the detector when flow is restarted. It is important to prevent contamination of shared stock reagents or solutions and to avoid storing them in an inappropriate vessels—e.g., acetonitrile, trifluoroacetic acid, concentrated acetic, and formic acids may leach UV-absorbing materials from certain plastic tubes or stoppers. Use glass or polypropylene where possible, and beware of black rubber disks (even if faced with PTFE) in solvent bottle stoppers. Always start a series of runs with a blank injection of HPLC solvent A; run some standards (e.g., glucagon and insulin, obtainable from Sigma) if peaks are broad or fail to elute (expect analytical peaks that are not overloaded to have a width at half height of ~30 sec with a gradient of 2% B/min). Ask a chromatography expert for advice if problems continue.

Correct fragmentation (in general, cleavage between Cys residues, without scrambling) is crucial to success. Depending on the known or expected amino acid sequence, digestion conditions are chosen that will cleave the polypeptide chain at least once between successive Cys residues. Acidic conditions (for cyanogen bromide, see Basic Protocol 2 and *UNIT 11.4*; for pepsin, see Basic Protocol 2) help to avoid disulfide scrambling. If neutral or basic conditions must be used (see *UNIT 11.1*), it is important to keep the pH as low as possible while preserving adequate enzyme activity; keep digestion times short by using a high protein concentration and enzyme/substrate ratio; and, if possible, operate in the presence of a thiol-specific reagent to trap any transient thiols. For recombinant CD23, for example, the tryptic digestion was performed at pH 7.0 in the presence of 10 mM iodoacetic acid (Rose et al., 1992). Note that commercially available asparaginyl-endopeptidase (Pierce), in contrast with Asp-N protease (*UNIT 11.1*), requires reducing conditions and so cannot generally be used for disulfide bond assignments (Rose et al., 1992). If saving time is more important than the amount of material consumed, perform several trial fragmentations in parallel and examine them analytically by reversed-phase HPLC.

Sometimes an isolated peptide component may not be susceptible to Edman degradation: it may be N-terminally blocked (e.g., by formyl or acetyl if it is the N-terminal peptide of the protein, or by a residue of pyroglutamic acid if a Gln residue liberated during proteolysis has cyclized). This problem is easily recognized by the fact that more components are found by mass spectrometry than by Edman degradation. It can usually be solved more easily by amino acid analysis of the isolated (reduced) component than by digestion with acyl amino acid

BSU IASST CENTRAL UNIVERSITY LIBRARY

hydrolase or pyroglutamyl amino peptidase to expose an amino group for a further attempt at Edman degradation. Alternatively, sophisticated mass spectrometric sequencing techniques may be employed; however, these are not available in the average biochemical laboratory.

Titration curve analysis

Avoid overfilling the sample trench with protein solution! Sample running freely on the gel surface will produce smears and curtains. If some sample overflows outside the trench accidentally, blot it immediately with filter paper.

If condensation is seen on the glass lid supporting the electrodes, it means that too much joule heat has been generated. This could be due to too high a power setting (i.e., 10 W); if this is the case, lower the setting. Most often, however, condensation comes from electrode filter paper strips being too wet with electrode solution; this results in solution flooding the gel surface in the first-dimension run. The remedy, at this point, would be to excise a wider area of gel in front of the electrode wicks when they are removed in the passage from the first gel to the second dimension (see Fig. 7.3.1B for this operation). Occasionally, the thermostatic water bath might not have been switched on, the coolant might not be recirculating, or the water bath might be jammed (in the Pharmacia Biotech instrument, press the button release to restart the thermostat).

If, upon staining, the protein titration curves appear too wavy or excessively broad, there are three things to check. First, inspect the contact between the platinum wire and the electrode strips. The platinum wire might have been damaged or broken, or might be dislodged from its groove. Integrity, good contact, and good alignment of the platinum wire are essential for good technique. Second, check the amount of salt in the protein (if enough liquid volume is available, one can use a conductivity meter). If too much salt is present, the high sample conductivity might generate too much heat in the trench and provoke convective flows or even partial protein denaturation due to heat shock or generation of strongly acidic and/or strongly basic ion boundaries. Remember that use of salts made of strong acids or bases is an anathema to all focusing techniques. Finally, check the protein concentration. If too much protein has been applied, the titration curves will be rather broad, since diffusion is proportional to the initial sample concentration.

Anticipated Results

Peptide mapping

Note that partial digestion products are usually present in varying amounts in both chemical and enzymatic cleavage procedures. For this reason, if the recombinant protein will be compared with the native protein, it is important to treat both samples concurrently so that they can be compared directly. The large number of fragments that result are easily resolved by HPLC, giving many points of comparison between the recombinant and native proteins. In the event that native protein is not available (the most likely circumstance), then the peptide fragments of the recombinant protein must be analyzed by additional techniques, including amino acid analysis, N-terminal sequence analysis, and/or mass spectrometry.

Analyzing peptide maps by high-resolution gel electrophoresis, on the other hand, may fail to resolve all of the fragments. While Tris-tricine gels are ideal for resolving peptides in fairly tight bands down to 2 kDa (*UNIT 10.1*), the error in mass measurement can be as high as 5%. Fragments close in size may not be resolved from each other by this method. In general, several digests may be required to get a good set of data to compare with the native protein. In the event that a native protein is not available for comparison, fragments can be transferred to PVDF membranes and analyzed by amino acid analysis or N-terminal sequence analysis.

Disulfide linkage analysis

Interpretion of the pairs of chromatograms (attempted reduction and control) is usually straightforward. If no disulfide bond is present, there should be essentially no difference in retention time between the peak present in the attempted reduction and in the control. If a disulfide bond is present, then either this bond is holding two small peptide chains together or it is holding a single polypeptide chain in a loop. In the former case, two peaks should be visible on reduction, whereas in the latter case all that is observed is a small shift to later retention time (opening of the loop exposes more of the peptide chain to the stationary phase).

If the results of the analytical reduction of a fraction are very clear-cut (small shift to later retention time, or splitting into two components), it may be sufficient to submit a portion of that fraction for Edman degradation, a portion for mass spectrometry, and a portion for in

BOULARD GENERAL INSTITUTE LIBRARY

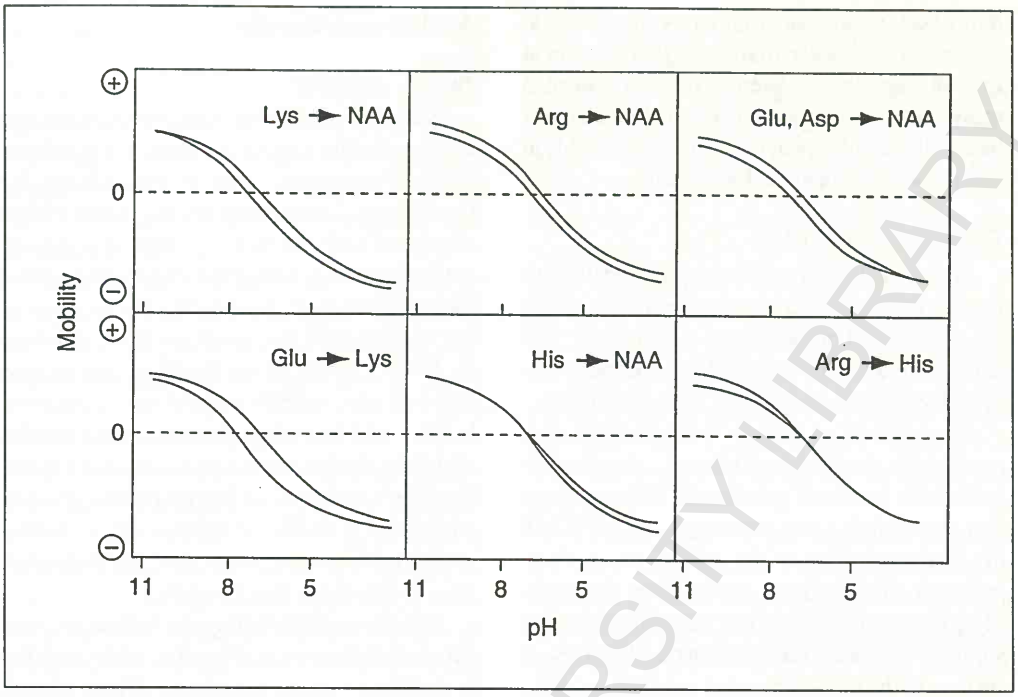


Figure 7.3.2 Theoretical titration curves of a wild-type protein and its genetic mutants in the presence of a given amino acid substitution. NAA: neutral amino acid; broken lines represent the zero-mobility plane (sample application trench); arrows indicate the direction of amino acid mutation. From Righetti et al. (1979).

situ reduction during mass spectrometry, thus avoiding the preparative reduction and HPLC step. Ask if the mass spectroscopist can perform in situ reduction: if this service is not available, then do preparative reduction, isolation, and analysis as described in Basic Protocol 2.

Titration curve analysis

Some examples of applications are presented in the figures. Figure 7.3.2 offers an example of differential titration curves obtained by mixing the presumptive mutant with the wild-type species. According to the various types of amino acid substitutions, different relative shapes will be detected, as shown. Note that only in the case of a replacement of arginine with a neutral amino acid (NAA) will the two curves run parallel all along the titration interval (typically not exceeding the pH 3 to 10 range).

Figure 7.3.3 gives a good example of affinity-titration curves obtained by incorporating a ligand into the gel matrix. In this case, a lectin from *Ricinus communis* was run in presence of increasing amounts of allyl-α-D-galactose grafted onto the gel matrix. This technique gives a unique depth of field, because not only does it permit measurement of the dissociation constant (K_d), but it also makes it possible to evaluate the variation in K_d with respect to pH (see Fig. 7.3.4).

Time Considerations

Peptide mapping. A peptide mapping experiment can be completed in ~2 days. The sample is generally digested overnight; analysis by either HPLC or high-resolution SDS-PAGE is then done on the second day. HPLC runs generally take ~2 hr each, including the pre-equilibration step for the column. Tris-tricine gels can be prepared ahead of time and stored overnight in a cold-box. Large gels generally require ~5 hr to run. Staining requires ~2 hr for Coomassie blue, whereas gels that will be silver stained should be washed with 50% methanol (four changes in 2 hr, or equilibration overnight) to remove substances that can interfere with the staining process. The silver staining process requires ~1.5 hr.

Disulfide linkage analysis. Each step (including dialysis) has been described in such a way as to make it possible to estimate how much time it will take. The time necessary for vacuum centrifugation and lyophilization steps depends somewhat on the scale, but samples containing ammonium bicarbonate must remain on the lyophilizer overnight or the salt

BU IASTRA UNIVERSITY LIBRARY

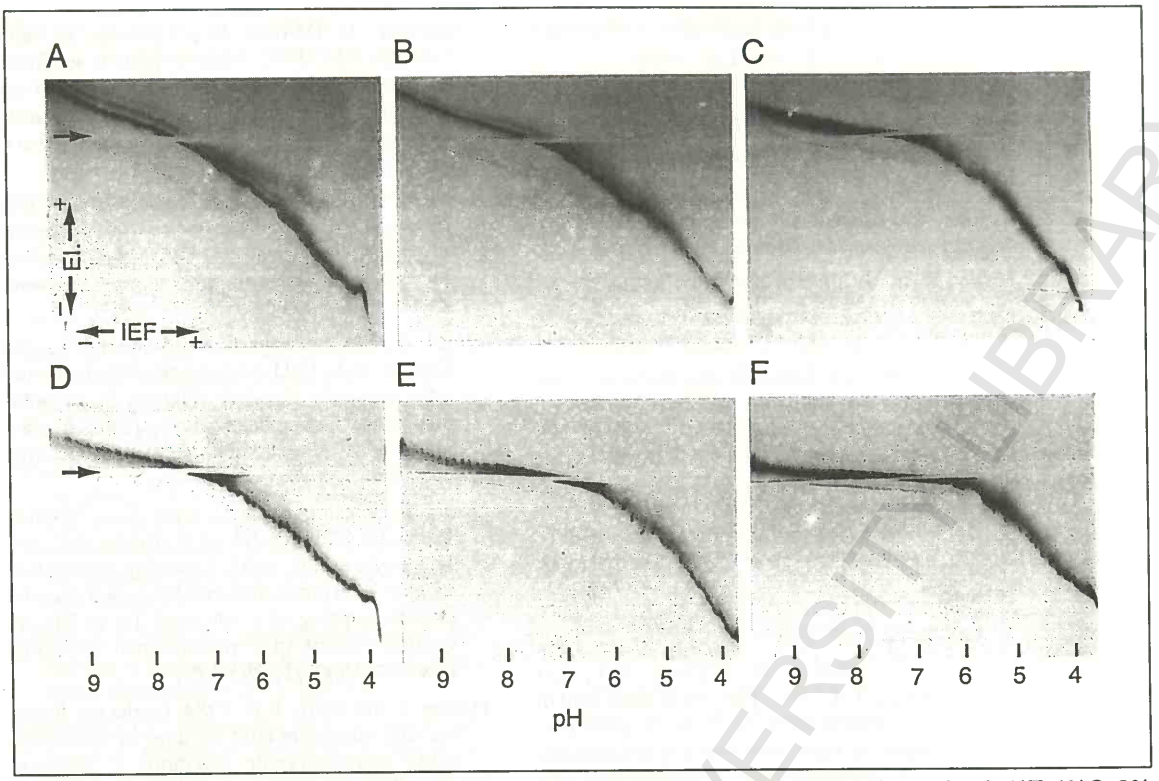


Figure 7.3.3 Affinity-titration curves of lectin from *Ricinus communis* seeds. The gel contained 6%T, 4%C, 2% Ampholine (pH 3.5 to 10), and 2 mM each Glu, Asp, Lys, and Arg. The amounts of allyl-α-D-galactose copolymerized in the gel matrix were: (**A**) 0 (no-ligand control); (**B**) 1×10^{-5} M; (**C**) 4×10^{-5} M; (**D**) 5×10^{-5} M; (**E**) 7×10^{-5} M; and (**F**) 10×10^{-5} M. The amount of protein loaded was 200 µg in 100 µl volume in all cases. The first- and second-dimension runs were 50 min at 10 W constant power and 20 min at 700 V constant voltage, respectively. The two arrows in (**A**) and (**D**) indicate the sample application trench (zero-mobility plane); the two double arrows with positive and negative symbols in (**D**) represent the direction and polarity of isoelectric focusing (IEF) and electrophoresis (El). From Ek et al. (1980).

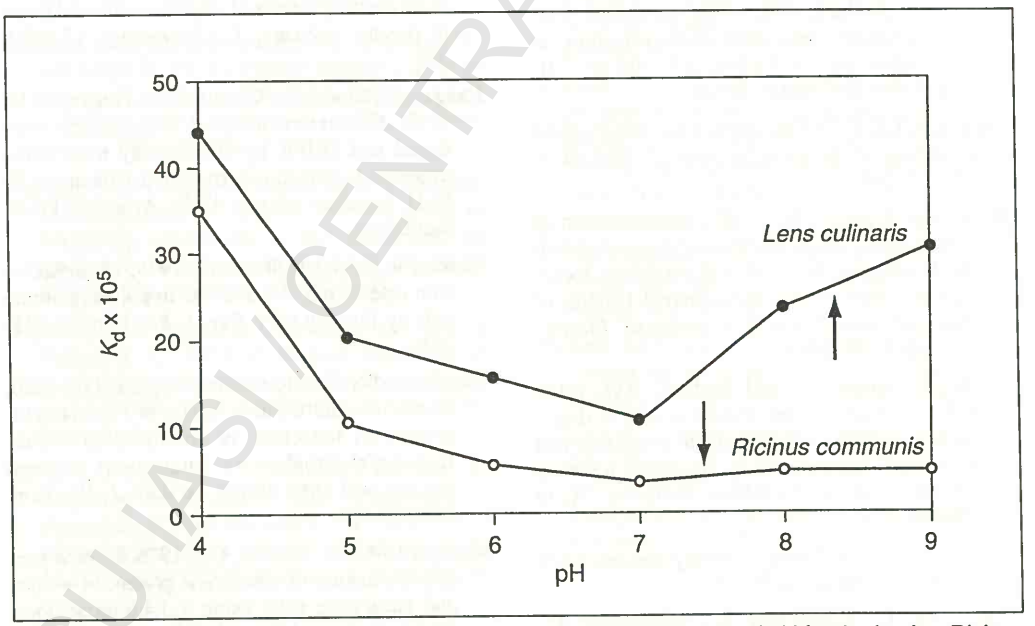


Figure 7.3.4 Variation of the dissociation constants (K_d) as a function of pH for the lectins *Ricinus communis* and *Lens culinaris*. These values were calculated from the retarded mobilities observed at different pH values in affinity-titration curves such as the one reported in Figure 7.3.3. The two vertical arrows indicate the pI of each lectin. From Ek et al. (1980).

BIBLIOTECA CENTRALE UNIVERSITÀ IASI

will not vaporize fully and some will remain with the protein. If blocking (steps 1 to 3), cyanogen bromide treatment, and enzymatic digestion must all be performed, the bonding pattern should be established within 2 weeks provided that an automatic injector is available for the HPLC.

Titration curve analysis. The titration curve analysis can be done in one day. Casting the gel requires ~1 hr. The first dimension electrophoresis takes ~1 hr, and the second dimension run requires ~30 min. Staining and destaining the gel can take ~1 to 2 hr.

Literature Cited

Allen, G. 1989. Specific cleavage of the protein in sequencing of proteins and peptides. *In* Laboratory Techniques in Biochemistry and Molecular Biology, 2nd ed. (R.H. Burdon and P.H. van Knippenberg, eds.) pp. 73-104. Elsevier Science Publishing, New York.

Arnaud, P., Gianazza, E., Righetti, P.G., and Fudenberg, H.H. 1980. The role of sialic acid in the microheterogeneity of α_1 acidic glycoprotein: Study by isoelectric focusing and titration curves. *In* Electrophoresis '79 (Radola, B.J., ed.) pp. 151-163. Walter de Gruyter, Berlin.

Bianchi-Bosisio, A., Loherlein, D., Snyder, R., and Righetti, P.G. 1980. Protein titration curves by isoelectric focusing-electrophoresis in highly porous (highly cross-linked) polyacrylamide gels. *J. Chromatogr.* 189:317-330.

Constans, J., Viau, M., Gouaillard, C., Bouisson, C., and Clerc, A. 1980. Binding affinities between VDBP and vitamin D_3,25-(OH)-D_3, 24-25-$(OH)_2$-D_3 and I-25-$(OH)_2$-D_3 studied by electrophoretic methods (PAGE-IEF combined IEF-electrophoresis) and print-immunofixation. *In* Electrophoresis '79 (Radola, B.J., ed.) pp. 701-710. Walter de Gruyter, Berlin.

Creighton, T.E. 1979. Electrophoretic analysis of the unfolding of proteins by urea. *J. Mol. Biol.* 129:235-264.

Ek, K. and Righetti, P.G. 1980. Determination of protein ligand dissociation constants and of their pH-dependence by combined isoelectric focusing electrophoresis (titration curves). Binding of phosphorylases A and B to glycogen. *Electrophoresis* 1:137-140.

Ek, K., Gianazza, E., and Righetti, P.G. 1980. Affino-titration curves: Determination of dissociation constants of lectin-sugar complexes and of their pH-dependence by isoelectric focusing electrophoresis. *Biochim. Biophys. Acta* 626:356-365.

Ellman, G.L. 1959. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82:70-77.

Fägerstam, L., Söderberg, L., Wahlström, L., Fredriksson, U.B., Plith, K., and Walldèn, E. 1983. Basic principles used in the selection of monobeads ion exchangers for the separation of biopolymers. *Prot. Biol. Fluids* 30:621-628.

Fernandez, J., DeMott, M., Atherton, D., and Mische, S.M. 1992. Internal protein sequence analysis: Enzymatic digestion for less than 10 µg of protein bound to polyvinylidene difluoride or nitrocellulose membranes. *Anal. Biochem.* 201:255-264.

Gianazza, E., Gelfi, C., and Righetti, P.G. 1980. Isoelectric focusing followed by electrophoresis of proteins for visualizing their titration curves by zymogram and immunofixation. *J. Biochem. Biophys. Methods* 3:65-75.

Glocker, M.O., Arbogast, B., Schreurs, J., and Deinzer, M.L. 1993. Assignment of the inter- and intra-molecular disulfide linkages in recombinant human macrophage colony-stimulating factor using fast atom bombardment mass spectroscopy. *Biochemistry* 32:482-488.

Harris, R.J., Murnane, A.A., Utter, S.L., Wagner, K.L., Cox, E.T., Polastri, G.D., Helder, J.C., and Sliwkowski, M.B. 1993. Assessing genetic heterogeneity in production cell lines: Detection by peptide mapping of a low level Tyr to Gln sequence variant in a recombinant antibody. *Bio/Technology* 11:1293-1297.

Henner, J. and Sitrin, R.D. 1984. Isoelectric focusing and electrophoretic titration of antibiotics using bioautographic detection. *J. Antibiot.* 37:1475-1478.

Hojrup, P. and Magnusson, S. 1987. Disulphide bridges of bovine Factor X. *Biochem. J.* 245:887-892.

Krishnamoorthy, R., Bianchi-Bosisio, A., Labie, D., and Righetti, P.G. 1978. Titration curves of liganded hemoglobins by combined isoelectric focusing electrophoresis. *FEBS Lett.* 94:319-323.

Jokl, V., Dolejsovà, J., and Matusovà, M. 1979. Zone electrophoresis of organic acids and bases in organic solvents. *J. Chromatogr.* 172:239-248.

Langen, H., Sander, B., Vilbois, F., and Lahm, H.-W. 1993. Characterization of the proteins c-kit ligand and DHFR by electrospray mass spectrometry. *In* Techniques in Protein Chemistry IV (R.H. Angeletti, ed.) pp. 47-54. Academic Press, San Diego.

Laskey, R.A. and Mills, A.D. 1975. Quantitative film detection of ^3H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* 56:335-341.

Lostanlen, D., Gacon, G., and Kaplan, J.C. 1980. Direct enzyme titration curve of NADH:cytochrome b_5 reductase by combined isoelectric focusing/electrophoresis. Interactions between enzyme and cytochrome b_5. *Eur. J. Biochem.* 112:179-183.

Markwell, M.A.K. and Fox, C.F. 1978. Surface-specific iodination of membrane protein of viruses and eucaryotic cells using 1,3,4,6-tetrachloro-3α,6α-diphenylglycoluril. *Biochemistry* 17:4807-4817.

BSU IASRI CENTRAL UNIVERSITY LIBRARY

Morris, H.R. and Pucci, P. 1985. A new method for rapid assignment of S-S bridges in proteins. *Biochem. Biophys. Res. Commun.* 126:1122-1128.

Niekamp, C.W., Hixson, H.F., and Laskowski, M. 1969. Peptide-bond hydrolysis equilibria in native proteins. Conversion of virgin into modified soybean trypsin inhibitor. *Biochemistry* 8:16-22.

Patterson, S.D. 1994. From electrophoretically separated protein to identification: Strategies for sequence and mass analysis. *Anal. Biochem.* 221:1-15.

Picard, B., Goullet, P., and Krishnamoorthy, R. 1987. A novel approach to study of the structural basis of enzyme polymorphism. *Biochem. J.* 241:877-881.

Proudfoot, A.I., Davies, G.D., Turcatti, G., and Wingfield, P.T. 1991. Assignment of the disulfide bridges of unglycosylated recombinant interleukin-5 expressed in recombinant *Escherichia coli* using proteolysis at low pH. *FEBS Lett.* 238:61-64.

Rabilloud, T. 1990. The mechanism of protein silver staining in polyacrylamide gels: A 10-year synthesis. *Electrophoresis* 11:785-794.

Randhawa, Z.I., Witkowska, H.E., Cone, J., Wilkins, J.A., Hughes, P., Yamanishi, K., Yasada, S., Masui, Y., Arthur, P., Kletke, C., Bitsch, F., and Shackleton, C.H.L. 1994. Incorporation of norleucine at methionine positions in recombinant human macrophage stimulating factor (M-CSF, 4-153) expressed in *Escherichia coli*: Structural analysis. *Biochemistry* 33:4352-4362.

Riesner, D., Henco, K., and Steger, G. 1991. Temperature gradient gel electrophoresis. *In* Advances in Electrophoresis (Chrambach, A., Dunn, M.J., and Radola, B.J., eds.) vol. 4, pp. 169-250. VCH Publishers, Weinheim and New York.

Righetti, P.G., Krishnamoorthy, R., Gianazza, E., and Labie, D. 1978a. Protein titration curves by combined isoelectric focusing-electrophoresis with hemoglobin mutants as models. *J. Chromatogr.* 166:455-460.

Righetti, P.G., Gacon, G., Gianazza, E., Lostanlen, D., and Kaplan, J.C. 1978b. Titration curves of interacting cytochrome b_5 and hemoglobin by isoelectric focusing electrophoresis. *Biochem. Biophys. Res. Commun.* 85:1575-1581.

Righetti, P.G., Menozzi, M., Gianazza, E., and Valentini, L. 1979. Protolytic equilibria of doxorubicin as determined by isoelectric focusing and electrophroetic titration curves. *FEBS Lett.* 101:51-55.

Rose, K., Turcatti, G., Graber, P., Pochon, S., Regamey, P.-O., Jansen, K.U., Magnenat, E., Aubonney, N., and Bonnefoy, J.-Y. 1992. Partial characterization of natural and recombinant human soluble CD23. *Biochem. J.* 286:819-824.

Rosengren, A., Bjellqvist, B., and Gasparic, V. 1977. A simple method of choosing optimum pH conditions for electrophoresis. *In* Electrofocusing and Isotachophoresis (Radola, B.J. and Graesslin, D., eds.) pp. 165-172. Walter de Gruyter, Berlin.

Rüchel, R. and Trost, M. 1981. A study of the structural conversions of two carboxyl proteinases employing electrophoresis across a pH gradient. *In* Electrophoresis '81 (Allen, R.C. and Arnaud, P., eds.) pp. 667-676. Walter de Gruyter, Berlin.

Sanger, F. 1959. The chemistry of insulin. *Science* 129:1340-1344.

Schrimsher, J.L., Rose, K., Simona, M.G., and Wingfield, P.T. 1987. Characterization of human and mouse granulocyte-macrophage-colony-stimulating factors derived from *Escherichia coli*. *Biochem. J.* 247:195-199.

Stahl, E. and Muller, J. 1981. pH-gradient-dunnschicht-chromatographie von benzodiazepinen. *J. Chromatogr.* 209:484-488.

Stults, J.T. 1995. Matrix-assisted laser desorption/ionization mass spectroscopy (MALDI-MS). *Curr. Opin. Structur. Biol.* 5:691-698.

Tate, M.E. 1981. Determination of ionization constants by paper electrophoresis. *Biochem. J.* 195:419-426.

Thannhauser, T.W., Konishi, Y., and Scheraga, H.A. 1984. Sensitive quantitative analysis of disulfide bonds in polypeptides and proteins. *Anal. Biochem.* 138:181-188.

Thornton, J.M. 1981. Disulfide bonds in globular proteins. *J. Mol. Biol.* 151:261-287.

Troungos, C., Krishnamoorthy, R., Elion, J., and Labie, D. 1982. Titration curves by combined isoelectric focusing-electrophoresis on a thin layer of agarose gel. 1982. *J. Chromatogr.* 250:73-79.

Valentini, L., Gianazza, E., and Righetti, P.G. 1980. pK determinations via pH-mobility curves obtained by isoelectric focusing electrophoresis. Theory and experimental verification. *J. Biochem. Biophys. Methods* 3:323-338.

Van Den Oetelaar, P.J. and Hoenders, H.J. 1987. Electrophoretic titration curve in 6 M urea of the bovine eye lens protein α-crystallin. *J. Chromatogr.* 411:507-509.

Vickers, M.F. and Robinson, P.A. 1983. Protein titration curves using modified cellulose acetate membranes. *J. Chromatogr.* 275:428-431.

Violand, B.N., Schlitter, M.R., Lawson, C.Q., Kane, J.K., Siegal, N.R., Smith, C.E., Kolodziej, E.W., and Duffin, E.L. 1994. Isolation of *Escherichia coli* synthesized eukaryotic proteins that contain ε-*N*-acetyllysine. *Protein Sci.* 3:1089-1097.

BCU SCIENCE CENTRAL LIBRARY

Wingfield, P., Mattaliano, R., MacDonald, R.H., Craig, S., Clore, G.M., Gronenborn, A.M., and Schmeissner, U. 1987. Recombinant-derived interleukin-1a stabilized against specific deamidation. *Protein Eng.* 1:413-417.

Wingfield, P., Benedict, R., Turcatti, G., Allet, B., Mermod, J.-J., DeLamarter, J., Simona, M.G., and Rose, K. 1988. Characterization of recombinant-derived granulocyte-colony stimulating factor (G-CSF). *Biochem. J.* 256:213-218.

Key Reference

Allen, G. 1989. See above.

Contains excellent information about common proteolytic agents.

Contributed by Nancy D. Denslow (peptide mapping)
University of Florida
Gainesville, Floria

Keith Rose (disulfide linkage analysis)
University Medical Center
Geneva, Switzerland

Pier Giorgio Righetti (two-dimensional titration curve analysis)
University of Milano
Milan, Italy

DCU IASI / CENTRAL UNIVERSITY LIBRARY

Transverse Urea-Gradient Gel Electrophoresis

Monitoring the cooperative unfolding transition induced when a protein is exposed to elevated temperature or a chemical denaturant is an important strategy for characterizing the conformational properties of a globular protein. This transition may be analyzed quantitatively by a variety of spectroscopic techniques, but a simpler alternative is urea-gradient gel electrophoresis. The pattern produced in the resulting gel can be used to estimate both the free energy change for unfolding and the rate of the unfolding transition. In addition, the technique can help identify either covalent or conformational heterogeneity in a protein sample. Because urea-gradient gel patterns are sensitive to several parameters, including hydrodynamic volume, net charge, and conformational stability, the technique can be particularly useful for comparing two forms of a protein, e.g., a natural form and the product of recombinant bacteria.

Transverse urea-gradient gels are slab gels prepared with a gradient of urea concentration perpendicular to the direction of electrophoresis (Creighton, 1979). A single sample is applied to the top of the gel, and as the protein is electrophoresed, molecules at different positions across the gel are exposed to different urea concentrations. Unfolding is detected by a decrease in electrophoretic mobility, reflecting the increased hydrodynamic volume of the unfolded protein.

The gels are prepared with the glass plates rotated 90° from the orientation used for electrophoresis, which requires minor modifications of most standard electrophoresis equipment. The major requirements are spacers to fit the sides of the glass plates as oriented for casting and some arrangement for holding the glass plates in this orientation. An arrangement devised for preparing gels for the Bio-Rad Mini-Protean II apparatus is illustrated in Figure 7.4.1. The spacers and a casting box for preparing multiple gels simultaneously are available from Aquebogue. The gels are prepared using a photoactivating polymerization catalyst to allow sufficient time to prepare the gradient. After the gels

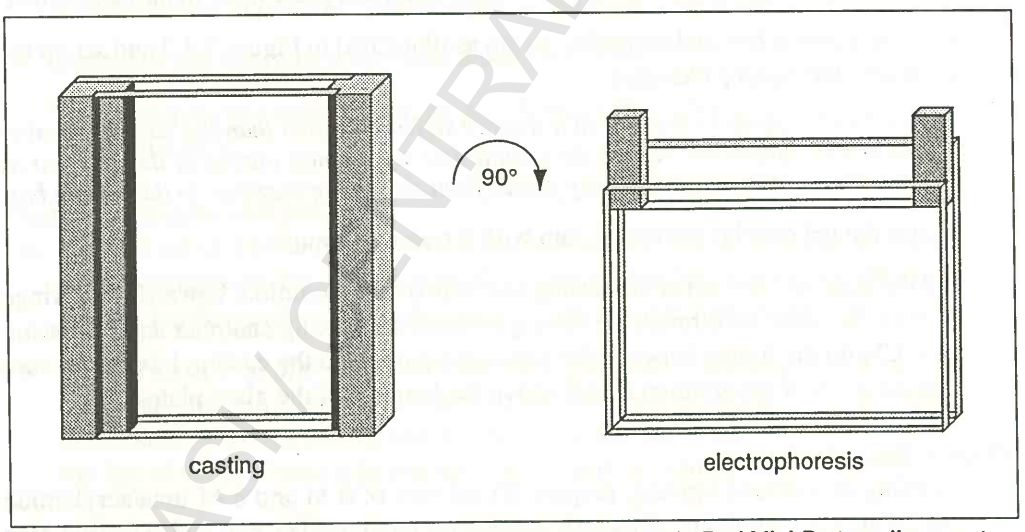


Figure 7.4.1 Spacers for preparing urea-gradient gels for the Bio-Rad Mini-Protean II apparatus. **(A)** A gel "sandwich" assembled and oriented for casting. The spacers are designed fit along the sides of the gel, as oriented for casting, and to keep the glass plates from sliding when placed in a casting box. **(B)** After the gel has polymerized, the gel is rotated 90°, the casting spacers are removed, and short spacers are placed on the sides of the gel to form a single sample well. The spacers are made to prepare gels 1 mm thick and can be constructed by machining single pieces of plastic or by glueing layers together. The standard glass plates for the Mini-Protean II apparatus are used. In the figure, the thicknesses of the spacers and glass plates have been exaggerated for clarity.

BYU HBL/CENTRAL UNIVERSITY LIBRARY

Copyright © 1996 by John Wiley & Sons, Inc.

have polymerized, the spacers are removed to expose the top and bottom of the gel, as oriented for electrophoresis. The gel is then preelectrophoresed and the sample is loaded. After electrophoresis, the gel is stained and analyzed.

The following instructions assume that the total volume of the gels is 40 ml for five 10-cm-wide gels in the Aquebogue casting box. The gels are prepared so that 2 cm on the left and right side (as oriented for electrophoresis) contain 0 and 8 M urea, respectively. Volumes should be adjusted appropriately for gels of other dimensions.

Materials

Gel overlay solution: 20% (v/v) ethanol/0.002% (w/v) bromophenol blue
Urea/acrylamide gel solutions with photopolymerization catalysts (see recipes)
10× electrophoresis buffer (see recipe)
Protein samples
Acidic *or* basic protein sample buffer (see recipes)
Coomassie blue R-250 stain solution (see recipe)
Gel rinse solution: 50% (v/v) methanol/7.5% (v/v) acetic acid
Gel destain solution: 5% (v/v) methanol/7.5% (v/v) acetic acid

Gel electrophoresis tank and glass plates (e.g., Bio-Rad)
Spacers for casting gels (Aquebogue)
Gel casting box (Aquebogue)
Three-channel peristaltic pump (e.g., ISCO Tris pump)
Light source to initiate photopolymerization of gels (e.g., two 15-W blue fluorescent lamps)
Air-displacement pipettor
Gel-loading tip
Plastic boxes

Set up apparatus for casting gradient gels

1. Assemble glass plates and spacers into sandwiches and place them in the casting box.
2. Connect casting box and peristaltic pump as illustrated in Figure 7.4.2 and set up the reservoir and mixing chamber.

   Position the pump so that it is at a slightly higher elevation than the mixing chamber and slightly below the inlet to the casting box to minimize mixing of the gradient as the solution of increasing density passes from the mixing chamber to the casting box.

3. Degas the gel overlay solution 5 min with a water aspirator.
4. Fill the bottom 2 to 4 cm of the casting box with overlay solution. Use a 20-ml syringe to draw the solution through the tubing between the mixing chamber and the casing box. Clamp the tubing between the peristaltic pump and the casting box. Make sure that the level of the solution is still above the bottoms of the glass plates.

Prepare the gel solutions

5. Working in subdued lighting, prepare 30 ml each of 0 M and 8 M urea/acrylamide gel solutions containing photopolymerization catalyst. Do not add TEMED (for riboflavin polymerization) or methylene blue (for methylene blue polymerization).

   Because of the time required to prepare the gradients, it is convenient to use a photoactivated polymerization catalyst, such as riboflavin or methylene blue. The gels can be cast under subdued lighting (complete darkness is not necessary), and then exposed to direct light to initiate polymerization.

   Appropriate pH and acrylamide concentrations of gel solutions must be chosen for each protein (see Critical Parameters).

ECU IASI GENERAL UNIVERSITY LIBRARY

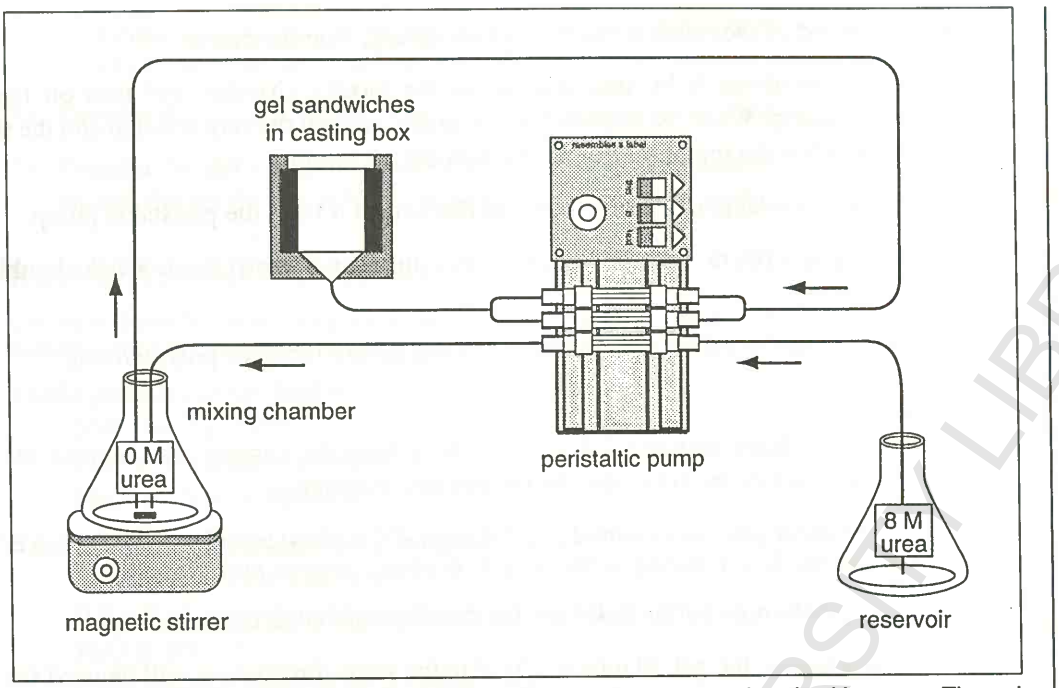


Figure 7.4.2 Preparation of urea-gradient gels using a three-channel peristaltic pump. The gel sandwiches are held in a clear plastic casting box with an inlet for the gel solution placed below the bottoms of the glass plates. Gel solutions containing 0 and 8 M urea are placed in the mixing chamber and reservoir respectively. As solution from the mixing chamber is pumped into the casting box using two channels of the pump, the 8 M urea solution is pumped into the mixing chamber using a single channel, leading to a linear increase in urea concentration in the mixing chamber. Before the gel solution is pumped into the casting box, the bottom of the box and the tubing between it and the mixing chamber are filled with gel overlay solution. The flow rate from the pump to the casting box should be ~2 ml/min.

Urea should be dissolved just prior to preparing the gels, because urea in solution breaks down to produce cyanate that can react with protein amino groups.

6. Degas the gel solutions 5 min with a water aspirator.

7. For photopolymerization, add 36 µl TEMED (for riboflavin) *or* 1.5 ml 2 mM methylene blue.

Cast the urea-gradient gels

8. Place 12 ml of 8 M urea gel solution in the reservoir. Use a 20-ml syringe to fill the tubing connecting the reservoir to the mixing chamber. Clamp the tubing in the peristaltic pump and temporarily place the outlet in the 8 M urea reservoir.

9. Place 8 ml of 0 M urea solution in the mixing chamber. Remove the clamp between the pump and the casting box. Turn on the peristaltic pump to begin pumping 0 M urea solution into the casting box at a rate of ~2 ml/min (1 ml/min per channel). When the last of the 0 M urea solution has just entered the tubing, stop the pump.

Be careful not to allow any air into the tubing.

10. Place 12 ml of 0 M urea solution in the mixing chamber and position the outlet from the 8 M urea solution in the mixing chamber. Turn on the magnetic stirrer and the peristaltic pump.

The concentration of urea in the flask will now increase linearly as the solution is pumped into the casting box.

BCU JAMES CENTRAL UNIVERSITY LIBRARY

11. When the last of the solution has entered the tubing, stop the pump.
12. Place the remaining 8 M urea solution in the mixing chamber and turn on the peristaltic pump. When the boundary between the ethanol overlay solution and the 0 M urea reaches the top of the gel plates, turn off the pump.
13. Clamp off the inlet to the casting box and disconnect it from the peristaltic pump.
14. Place the gels ~10 cm from the light source to initiate polymerization, which should take ~30 min.
15. Promptly rinse the gel solution out of the tubing before it begins polymerizing.

Set up the gel

16. After the gels have polymerized, remove them from the casting box, remove the spacers, and mount the gel in the electrophoresis chamber.

    Urea-gradient gels can be stored 1 to 2 days at 4°C without significant dissipation of the gradient. Before storing, wrap the gels in plastic wrap to prevent drying.

17. Dilute electophoresis buffer to 1× and fill the electrophoresis tank.
18. Preelectrophorese the gel 30 min at 100 V in the same direction as will be used for electrophoresis.

    For negatively charged proteins, electrophorese with the positive electrode at the bottom of the gel; for positively charged proteins, electrophorese with the negative electrode at the bottom of the gel.

    These conditions are usually adequate to remove reactive compounds resulting from the breakdown of urea or the polymerization process in a gel that is 1 mm thick and 10 cm wide.

Prepare and electrophorese samples

19. Prepare the sample for electrophoresis by diluting ~50 µg of protein to a total volume of 60 µl. Add 15 µl sample buffer appropriate for acidic or basic protein.

    Small samples of unfolded proteins in 8 M urea can be prepared by dissolving 48 mg solid urea in 63 µl aqueous solution (including the protein) to yield a total of 100 µl. It is not necessary to add glycerol to samples that contain 8 M urea.

20. Apply the sample evenly across the top of the gel using an air-displacement pipettor with a gel-loading tip.
21. Electrophorese the sample.

    The voltage and duration of electrophoresis appropriate for a given protein and set of electrophoresis conditions must be determined empirically, but 50 to 100 V for 3 to 4 hr is a good starting point.

22. At the end of the run, turn off the power supply and disconnect the electrodes. Remove the gel from the apparatus and remove the plates. Mark the orientation of the gel by cutting off one corner.

Stain and destain the gel

23. Place the gel in a plastic box containing ~50 ml Coomassie blue R-250 stain solution. Agitate gently ≥2 hr at room temperature.

    The high acid concentration of the recommended stain makes this solution particularly effective for fixing smaller proteins that are not easily precipitated by milder staining conditions. For some applications, staining solutions commonly used for SDS gel

ECU IAS / CENTRAL UNIVERSITY LIBRARY

electrophoresis—e.g., 0.1% (w/v) Coomassie blue R-250/50% (v/v) methanol/7.5% (v/v) acetic acid—may also be suitable.

24. Transfer the gel to a clean plastic box containing gel rinse solution and rinse ≤5 min.

25. Transfer the gel to a plastic box containing gel destain solution. Agitate gently and change the gel destain solution as necessary to obtain a nearly clear background.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocols steps. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Acidic protein sample buffer

50% (w/v) glycerol
0.01% (w/v) bromophenol blue
Store at room temperature

Acrylamide/bisacrylamide stock solution

30 g acrylamide (30% final)
0.8 g *N,N′*-methylene bisacrylamide (0.8% final)
H_2O to 100 ml
Filter through paper or an 0.45-µm filter
Store at 4°C in the dark

CAUTION: *Acrylamide and bisacrylamide are toxic. The powders should be handled in a fume hood, and both the powder and solutions should be handled wearing gloves.*

Basic protein sample buffer

50% (w/v) glycerol
0.2% (w/v) methyl green
Store at room temperature

Coomassie blue R-250 stain solution

0.5 g Coomassie blue R-250 dissolved in ~400 ml H_2O (1 mg/ml final)
50 g trichloroacetic acid (0.61 M final)
50 g 5-sulfosalicylic acid (0.39 M final)
H_2O to 500 ml
Store at room temperature

CAUTION: *Trichloroacetic acid and 5-sulfosalicylic acid are very caustic. Wear gloves when handling the solids or stain solution.*

Electrophoresis buffer, 10×

Buffer solutions that have been used successfully for urea-gradient gels include:
0.5 M acetate-Tris, pH 4.0: 0.5 M acetic acid adjusted to pH 4.0 with Tris base
0.5 M imidazole/0.5 M MOPS, pH ~7.0
0.5 M Tris-acetate, pH 8.0: 0.5 M Tris base adjusted to pH 8.0 with acetic acid
0.5 M Tris/0.5 M Bicine, pH ~8.4
Store all buffer solutions at room temperature

Urea/acrylamide gel solutions containing photopolymerization catalysts

0 M urea/15% acrylamide gel solution for riboflavin-polymerized gels
15 ml acrylamide/bisacrylamide stock solution (see recipe)
3 ml 10× electrophoresis buffer (see recipe; 1× final)
3.75 ml 0.04 mg/ml riboflavin (5 µg/ml final)
8.25 ml H_2O

Immediately before casting gels, degas gel solution, then add 36 µl TEMED.

continued

BOSTON CENTRAL UNIVERSITY LIBRARY

8 M urea/11% acrylamide gel solution for riboflavin-polymerized gels

14.4 g solid urea
11 ml acrylamide/bisacrylamide stock solution (see recipe)
3 ml 10× electrophoresis buffer (see recipe; 1× final)
3.75 ml 0.04 mg/ml riboflavin (5 µg/ml final)
1.2 ml H_2O

Immediately before casting gels, degas gel solutio, then add 36 µl TEMED.

0 M urea/11% acrylamide gel solution for methylene blue–polymerized gels

11 ml acrylamide/bisacrylamide stock solution (see recipe)
3 ml 10× electrophoresis buffer (see recipe; 1× final)
1.5 ml 20 mM sodium toluene sulfinate (1 mM final)
1.5 ml 1 mM diphenyliodonium chloride (50 µM final)
11.5 ml H_2O

Immediately before casting gels, degas solution and add 1.5 ml 2 mM methylene blue (100 µM final).

8 M urea/7% acrylamide gel solution for methylene blue–polymerized gels

14.4 g solid urea
7 ml acrylamide/bisacrylamide stock solution (see recipe)
3 ml 10× electrophoresis buffer (see recipe; 1× final)
1.5 ml 20 mM sodium toluene sulfinate (1 mM final)
1.5 ml 1 mM diphenyliodonium chloride (50 µM final)
4.4 ml H_2O

Immediately before casting gels, degas solution and add 1.5 ml 2 mM methylene blue (100 µM final).

COMMENTARY

Background Information

One of the most general properties of folded globular proteins is the tendency to undergo a cooperative unfolding transition when exposed to elevated temperature or to chaotropic agents such as urea or guanidinium chloride (Tanford, 1968). The unfolding transition for a particular protein under defined conditions is characterized by several thermodynamic properties, including the free energy change for unfolding, the enthalpy change, the entropy change, and, even more fundamentally, whether or not the transition is sufficiently cooperative that only fully folded and fully unfolded molecules are present at equilibrium. The free energy change for unfolding, ΔG_u, is a particularly important parameter that represents the net stability of the folded conformation relative to the unfolded state and is widely used for comparing the stabilities of related proteins. Unfolding-refolding transitions are also characterized by their kinetic properties, including their rates and the accumulation of any transient intermediates.

The thermodynamics and kinetics of protein unfolding and refolding reactions have been studied extensively using a variety of different biophysical methods, including spectroscopic and calorimetric techniques. These studies have led to important advances in understanding the physical interactions that stabilize folded proteins and the mechanisms by which these structures form (Schellman, 1987; Privalov, 1989; Kim and Baldwin, 1990; Creighton, 1992; Matthews, 1993). Although spectroscopic techniques and calorimetry have been mastered in a growing number of laboratories specializing in protein stability and folding, these measurements are quite technically demanding and require equipment and expertise that is not always found in typical protein biochemistry laboratories.

Urea-gradient gel electrophoresis was developed by T.E. Creighton (1979) as a simple alternative to spectroscopic methods for studying protein folding and unfolding. In this procedure, a single sample is applied across the top of a polyacrylamide slab gel containing a horizontal gradient of urea concentration. As the sample is electrophoresed through the gel, molecules at different positions across the gel are exposed to increasing urea concentrations.

ECU IASH CENTRAL UNIVERSITY LIBRARY

Those molecules that unfold have a larger hydrodynamic volume than the folded protein molecules and therefore migrate more slowly. After electrophoresis, the stained gel produces a graphic representation of the unfolding transition. As discussed later (see Anticipated Results), the pattern generated can provide information about whether or not the protein undergoes a urea-induced unfolding transition, information about the presence of significantly populated intermediate states in this transition, and semiquantitative information about the net stability of the folded protein and the rates of interconversion between the native and unfolded states.

Although the data obtained from urea-gradient gels are not generally as quantitative as those derived from spectroscopic or calorimetric measurements, the electrophoretic method offers a number of advantages for many applications. In addition to being relatively simple to implement, urea-gradient gel electrophoresis requires only small amounts of protein—a 50-μg sample is sufficient for detection with Coomassie blue, and even less can be used with more sensitive detection methods (*UNIT 10.5 & 10.6*). Furthermore, heterogeneous samples may often be used, because different components are likely to be separated during electrophoresis (Hollecker and Creighton, 1982). Indeed, urea-gradient electrophoresis is a particularly sensitive method for detecting heterogeneity, because molecules that differ in their electrophoretic mobilities in either the native or unfolded form can be distinguished, as can molecules that differ only with respect to the urea concentration at which they unfold.

In addition to being convenient, urea-gradient gels can often provide information about protein-folding transitions that is complementary to the information obtained from more conventional biophysical methods. Most spectroscopic techniques—e.g., UV absorbance (*UNIT 7.1*), fluorescence, and circular dichroism—measure only the averaged properties of a sample. As a consequence, it is often difficult to determine if, for instance, the signal measured at an intermediate urea concentration arises exclusively from a mixture of fully unfolded and fully native molecules or if a partially folded species may be present. Separation techniques such as gel electrophoresis can sometimes help resolve such ambiguities, particularly if different conformational states interconvert relatively slowly on the time scale of the separation.

The most useful applications of urea-gradient gels are often those where a qualitative comparison of different proteins is required, as in comparing the protein produced in recombinant bacteria with that obtained from a natural source. If the recombinant-produced protein has the same covalent structure and has folded to the same conformation as that of the natural protein, the two samples should display identical electrophoresis patterns. If, on the other hand, the two proteins differ in their covalent structure, this might be reflected in differences in the mobilities of either the native or unfolded forms, particularly if there are differences in net charge. If the conformations of the folded proteins are different, this may lead to differences in the mobilities of the native form or in the urea concentration at which the proteins unfold. Other applications include screening protein variants for differences in folding thermodynamics and kinetics (Klemm et al., 1991; Creighton and Shortle, 1994) and carrying out preliminary experiments to define conditions for more detailed studies using spectroscopic methods.

Critical Parameters

Because the behavior of a protein during urea-gradient gel electrophoresis depends on several properties of the particular protein, including its hydrodynamic volume, net charge, and conformational stability, there are more parameters that must be considered than for SDS gel electrophoresis, for instance. These parameters include pH, direction of electrophoresis, gel composition, and duration of electrophoresis.

pH

Because electrophoresis depends on the intrinsic net charge of the protein, the pH must be sufficiently far from the isoelectric point to maintain a significant net charge, usually 5 to 10 charge units (the charge of a proton or electron). For proteins with extremely high or low isoelectric points, it will probably be necessary to work below or above the isoelectric point, respectively, but proteins with isoelectric points near neutrality can be electrophoresed under either acidic or basic conditions. Although it may be desirable, in general, to use a pH close to physiological, in some cases it may be necessary to use more extreme conditions. For instance, it may be necessary to use a higher or lower pH to maintain a large enough net charge on the protein for electrophoresis or to

BYU CENTRAL UNIVERSITY LIBRARY

keep the protein soluble, because many proteins are least soluble close to their isoelectric points. In addition, pH can be used as a means to manipulate the stability of the folded protein. Many proteins are too stable to be unfolded in 8 M urea at neutral pH, but can be unfolded by urea under alkaline or acidic conditions.

Several different buffers have been used successfully for urea-gradient gels (see Reagents and Solutions). Ideally, the ionic strength of the buffer solution should be as low as possible to minimize the heat generated during electrophoresis. Meeting this condition can be facilitated by using buffers in which both the anionic and cationic component provide buffering capacity; several buffer systems of this type have been described by McLellan (1982). The direction of electrophoresis is determined by the pH chosen. If the pH is above the isoelectric point, the protein will have a net negative charge and will migrate towards the anode in the electrophoresis tank, as in SDS–polyacrylamide gel electrophoresis. If, on the other hand, a pH below the isoelectric point is used, the orientation of the electrodes must be reversed.

Gel composition

As in other forms of electrophoresis, the gel composition must be chosen to match the hydrodynamic volume of the protein of interest, with larger proteins requiring lower acrylamide concentrations. The sieving properties of the gel and, therefore, the choice of acrylamide concentration also depend on the presence of urea in the gel and the method of polymerization used. For reasons that have not been fully characterized, urea in a polyacrylamide gel decreases the electrophoretic mobility, so that even in the absence of an unfolding transition there will be a progressive reduction of mobility across a urea-gradient gel. This can be largely compensated for by incorporating an acrylamide gradient so that there is a lower acrylamide concentration at the high-urea side of the gel. In his original protocol, Creighton (1979) used a 15% to 11% (w/v) acrylamide gradient (photopolymerized with riboflavin) superimposed on the 0 to 8 M urea gradient, and these conditions have worked well for a variety of monomeric proteins, including bovine serum albumin, with M_r = 66,000. Larger proteins, however, are likely to require lower acrylamide concentrations.

Gel sieving is also greatly influenced by the conditions used for polymerization. The two catalysts that have been most widely used to polymerize acrylamide gels are ammonium persulfate and riboflavin, the latter used for photoactivation. Both are used in conjunction with *N,N,N′,N′*-tetramethylethylenediamine (TEMED). Photopolymerization is particularly convenient for preparing gradient gels, because the timing of polymerization can be easily controlled. A major drawback to the use of riboflavin, however, is that polymerization is not as complete as when ammonium persulfate is used. Although this is not a serious difficulty for gels prepared at slightly acidic pH, at neutral or alkaline pH the gels may be very soft or the solution may not gel at all. The efficiency of photopolymerization can be increased by adding a small amount of ammonium persulfate, but it can be very tricky to find a concentration that will stimulate photopolymerization without also causing early polymerization. More recently, a procedure for photopolymerizing gels using methylene blue has been described, and it appears that this catalyst leads to much more efficient polymerization than riboflavin does (Lyubimova et al., 1993). Urea-gradient gels containing 11% to 7% (w/v) acrylamide and polymerized with methylene blue have been reported to have sieving properties similar to those of 15% to 11% (w/v) acrylamide gels polymerized with riboflavin (Creighton and Shortle, 1994).

Sample size

For urea-gradient gels ~10 cm wide, a sample volume of 75 µl is appropriate. This is large enough to apply evenly across the top of the gel and small enough to give reasonably narrow bands. It should be possible, in principle, to use discontinuous buffer systems with a stacking gel, but this does not appear to be necessary. Even in the absence of a discontinuous buffer system, there is usually some stacking effect when a macromolecular sample enters a gel, because the first molecules to enter the gel are retarded and accumulate near the top of the gel while the rest of the molecules enter the gel. The difference between the mobilities of molecules above and within the gel can be enhanced by making the ionic strength of the sample lower than that of the electrophoresis buffer.

Duration of electrophoresis

The time required for electrophoretic separation depends on the voltage applied to the gel, and this parameter can be manipulated according to the type of information desired. As discussed later (see Anticipated Results), the elec-

ECU IASTATE UNIVERSITY LIBRARY

trophoresis pattern generated depends to a large degree on whether the interconversion between native and unfolded states is rapid on the time scale of the electrophoresis. In particular, if the two forms interconvert many times during the electrophoresis, then a smooth, continuous band of protein should be generated, and the position of the observed transition can be used to estimate the thermodynamic stability of the protein. On the other hand, if the two forms interconvert slowly, a discontinuous or blurred band will be generated, and under favorable conditions it is possible to estimate the rate of the transition from the appearance of the bands and the duration of electrophoresis. Significant separations have been observed after as little as 8 min of electrophoresis (Creighton, 1980), but times as long as 10 hr have been used to generate smooth patterns for relatively slow transitions (Klemm et al., 1991).

An additional parameter that can be manipulated to advantage is the presence of other molecules or ions in the gel that influence the unfolding transition by binding selectively to either the native or unfolded forms. For instance, the influence of a divalent cation can be examined by comparing the patterns observed when the ion is present in the electrophoresis buffers with the patterns obtained when EDTA is added to the buffers. This approach is limited, of course, to compounds that do not interfere with either gel polymerization or electrophoresis. Thiol reagents, for instance, react with free radicals generated during polymerization, and high salt concentrations lead to very high temperatures during electrophoresis.

Troubleshooting

The difficulties most likely to arise in using urea-gradient gel electrophoresis are the formation of protein aggregates or covalent modifications either before or during the electrophoresis. Aggregation is likely to lead to smearing of the protein bands or a complete failure of the protein to enter the gel. It may be possible to minimize aggregation by choosing a different pH or ionic strength for electrophoresis or by decreasing the concentration of protein applied to the gel. Covalent modifications are likely to cause multiple protein bands, particularly if they lead to changes in net charge. Such modifications may arise from residual free radicals from the polymerization reaction or from cyanates generated by hydrolysis of urea. To minimize these reactions, the urea solutions should be prepared immediately before casting the gels, and the gels should be pre-electrophoresed before applying the sample.

Anticipated Results

The simplest patterns generated by urea-gradient gel electrophoresis arise when the native (*N*) and unfolded (*U*) forms of the protein are in rapid equilibrium and there are no significant populations of partially folded intermediates. In these cases, a simple sigmoidal band is produced (see Fig. 7.4.3).

If the two forms are in rapid equilibrium on the time scale of the electrophoresis, the band should remain sharp through the transition region, and the same pattern should be obtained whether native or urea-unfolded protein is applied to the gel. Provided that only the native

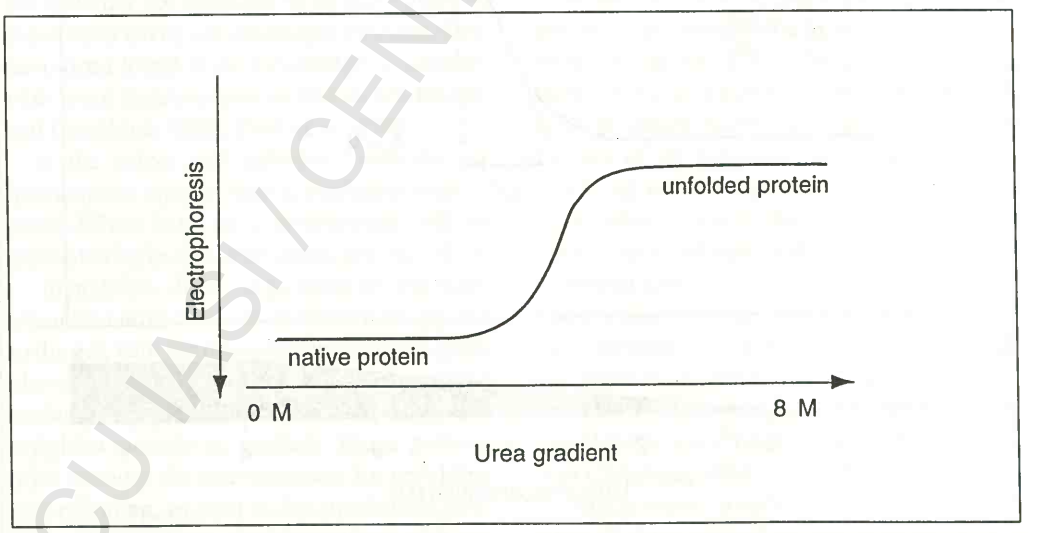


Figure 7.4.3 Diagram of an unfolding curve produced by urea-gradient gel electrophoresis of a protein undergoing rapid interconversion between the folded and unfolded states.

B. C. U. „M. EMINESCU" IAŞI
CHITIE
BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

and unfolded forms are significantly populated, the observed mobility at the intermediate urea concentrations will reflect the relative concentrations of N and U ([N], [U]), and the pattern can be interpreted as a plot of the fraction of protein unfolded versus urea concentration.

Although urea-gradient gels are used primarily for qualitative analysis of protein unfolding transitions, they may also be used to obtain an estimate of the conformational stability of a protein, provided that the conditions described above are met. The free energy change for unfolding is given by

$$\Delta G_u = -RT \ln \frac{[U]}{[N]}$$

where R is the gas constant and T is the absolute temperature. The value of ΔG_u will depend on the urea concentration, becoming more negative at higher concentrations. Because ΔG_u can only be measured at urea concentrations where N and U are both detectable, i.e., in the transition zone, it is usually necessary to extrapolate values measured near the middle of the transition to 0 M urea in order to estimate the stability under physiological conditions. Numerous studies of protein unfolding transitions have shown that the dependence of ΔG_u on denaturant concentration can usually be well described by a simple linear relationship,

$$\Delta G_u = m \cdot (C - C_m)$$

where C_m is the urea concentration at the transition midpoint (i.e., where $\Delta G_u = 0$, and the average mobility is midway between that of the native and unfolded forms), C is any other urea concentration and m is an empirical parameter (Schellman, 1978; Pace et al., 1989).

Hollecker and Creighton (1982) have shown that m is proportional to the derivative of the fraction unfolded (f_u) with respect to urea concentration evaluated at the midpoint according to the relationship:

$$m = -4RT \frac{df_u\,(C_m)}{dC}$$

This derivative can be estimated from urea-gradient gel patterns by drawing a line tangent to the gel band at the midpoint and measuring the change in urea concentration, ΔC, associated with the extrapolated change in f_u from 0 to 1. This procedure is illustrated in Figure 7.4.4.

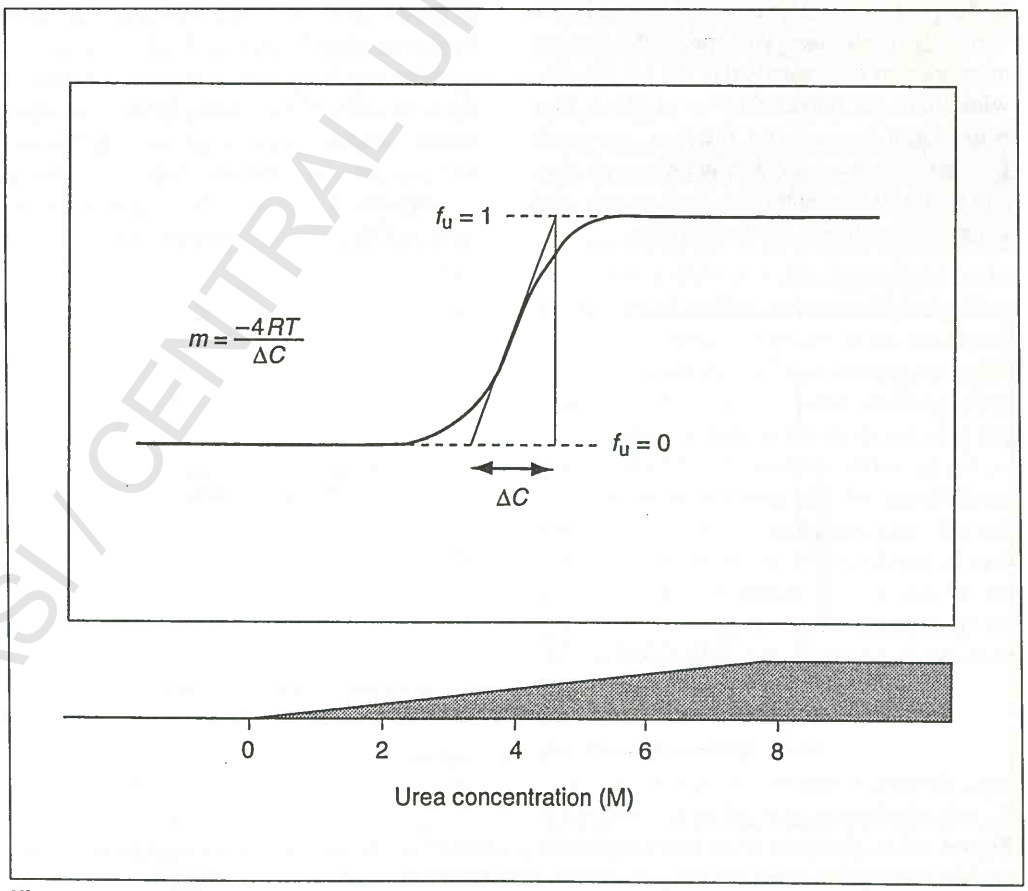


Figure 7.4.4 Estimating the parameter m from a urea-gradient gel. See text for details.

ECU IASI / CENTRAL UNIVERSITY LIBRARY

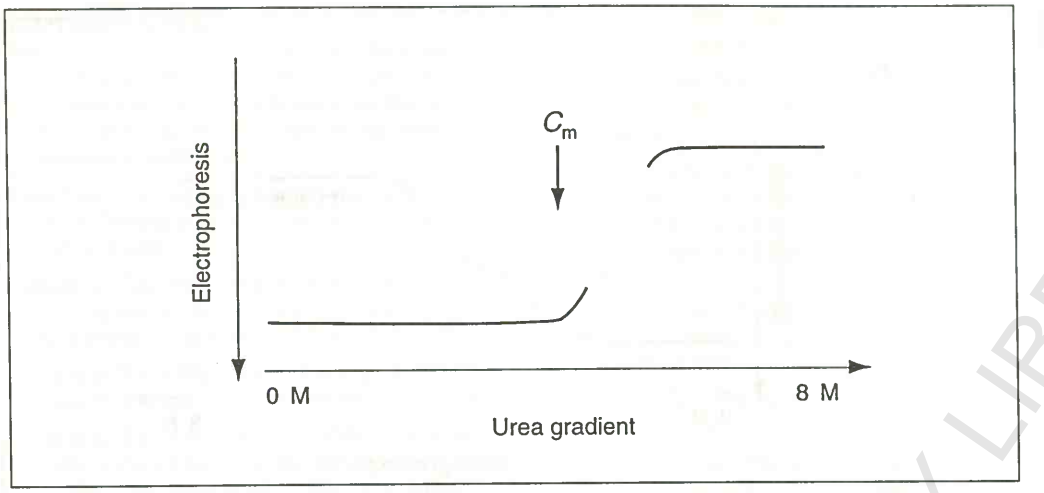


Figure 7.4.5 Urea-gradient gel pattern produced by a protein that slowly interconverts between the native and unfolded states when native protein is applied to the gel.

The derivative is simply $1/\Delta C$, and m is calculated according to:

$$m = \frac{-4RT}{\Delta C}$$

From the value of m and C_m, the value of ΔG_u can be calculated for any other urea concentration, including 0 M, where $\Delta G_u = -mC_m$ (Hollecker and Creighton, 1982; Klemm et al., 1991).

Quantitating urea-gradient gel results in this manner requires knowing the urea concentration at different positions across the gel, which can be estimated from the dimensions of the gel and from the known volumes of the linear gradient and the two bordering segments containing 0 and 8 M urea. Although values of C_m, m, and ΔG_u estimated from urea-gradient gels are certainly not expected to be as accurate as those obtained by spectroscopic methods, they have been found to be surprisingly consistent with more rigorous measurements (Hollecker and Creighton, 1982; Pace et al., 1989).

If the native and unfolded forms do not interconvert rapidly during electrophoresis, a more diffuse band or a discontinuity will be generated in the transition zone (see Fig. 7.4.5).

In addition, different patterns are expected when the native and unfolded forms are applied to the gel, with the transition appearing slightly above the equilibrium midpoint (C_m) when native protein is applied and below C_m when unfolded protein is applied. These patterns arise because the rate constants for unfolding and refolding, as well as the equilibrium constant, vary with urea concentration, with unfolding rates increasing and folding rates decreasing at higher urea concentrations. When native protein is applied to the gel, it remains in its initial state at low urea concentrations both because this state is thermodynamically favored and because unfolding is very slow. At urea concentrations near the equilibrium midpoint, the native conformation does not predominate thermodynamically, but at least a fraction of the molecules remain in this state simply because they do not have time to unfold during the time of electrophoresis. At urea concentrations somewhat above the midpoint, the unfolding rate is fast enough that most or all of the molecules unfold during the electrophoresis. When unfolded protein is applied to the gel, the opposite effect is seen, with a fraction of the molecules remaining unfolded even at urea concentrations below the equilibrium midpoint.

When the unfolding and refolding rates are not fast compared to the time of electrophoresis, the patterns cannot be used to estimate thermodynamic parameters, but it may be possible to obtain some information about the kinetics of the transition. If the half-time for unfolding (or folding) at the midpoint is approximately equal to the time of electrophoresis, a smeared band will be generated in the transition zone, and slower transitions will give rise to a discontinuity, as described above. The rate constants can be estimated by comparing the observed patterns with those predicted by numerical simulations, as discussed in greater detail elsewhere (Creighton, 1979; Goldenberg and Creighton, 1984; Goldenberg, 1989).

Much more complicated gel patterns can arise if there are additional species present either at equilibrium or accumulating as transient kinetic intermediates. If there is a partially

EDINBURGH UNIVERSITY LIBRARY

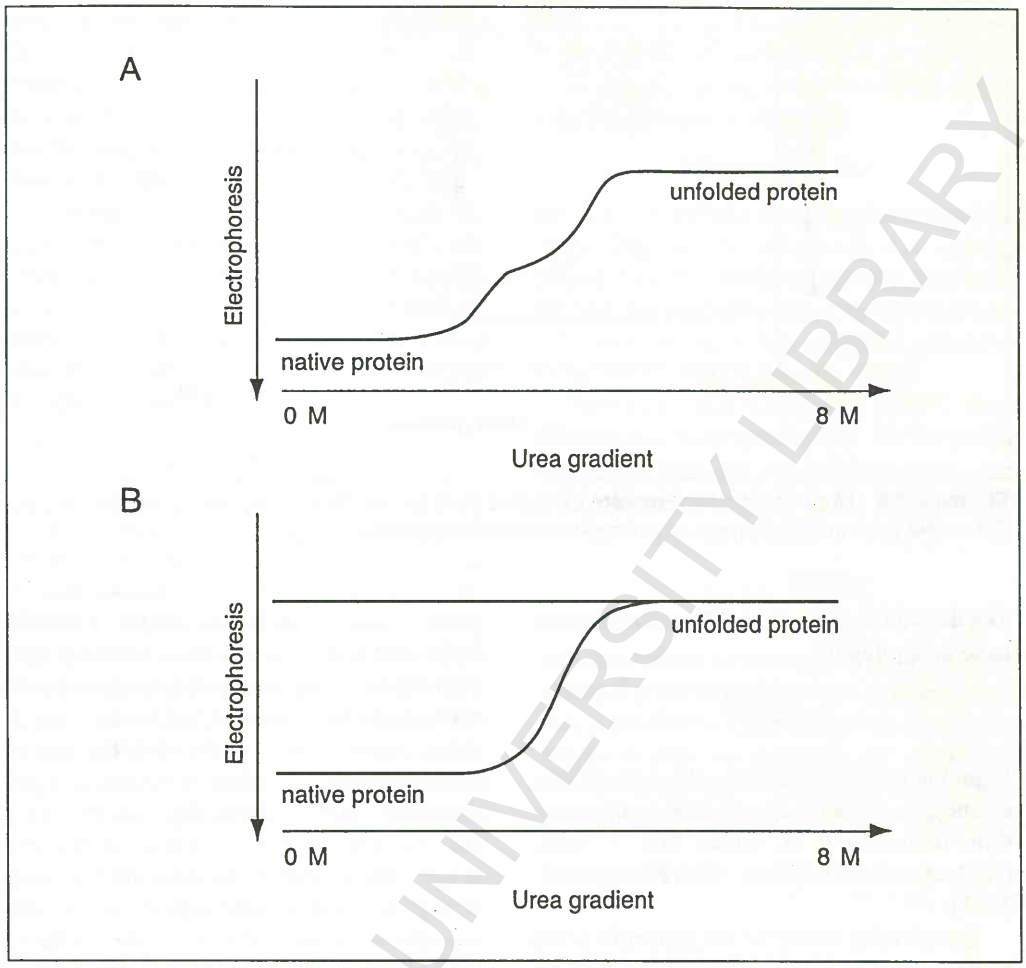


Figure 7.4.6 Urea-gradient gel patterns produced by proteins with additional species. **(A)** A protein with a partially unfolded species that predominates at intermediate urea concentrations. **(B)** A protein with two unfolded species, one of which refolds very slowly.

unfolded species that predominates at intermediate urea concentrations, a two-step transition may be observed (see Fig. 7.4.6A).

A frequently observed type of conformational heterogeneity is one in which there are multiple unfolded forms that refold at different rates. Multiple unfolded forms can arise from isomerization of prolyl peptide bonds (Brandts et al., 1975) and can give rise to the pattern depicted in Figure 7.4.6B when unfolded protein is applied to the gel (Creighton, 1980). This pattern is generated because some of the unfolded molecules can rapidly refold and establish an equilibrium with the native protein, but others do not have time to refold during the electrophoresis and remain unfolded even at urea concentrations at which the native protein is thermodynamically preferred. At the lowest urea concentrations, some of the slowly refolding form may have time to fold to the native form, giving rise to a blurred or discontinuous transition. Also, the slowly folding form may equilibrate with partially folded forms, giving rise to a downward curvature in the upper band. When native protein is applied to the gel, only the lower, sigmoidal, band is expected, because the slowly refolding form is generated relatively slowly after the protein unfolds. Other complexities in the patterns may arise from multiple folded forms, intermediates that convert slowly with the native or unfolded protein, or oligomeric proteins that undergo dissociation either prior to or in concert with unfolding.

Time Considerations

Most urea-gradient gel electrophoresis experiments can easily be completed in 1 day. A few hours should be allocated for preparing the gels, and electrophoresis may take as little as 20 min or as long as several hours. Once prepared, the gels can be kept at least 2 days at 4°C without apparent changes in their properties.

ECU IASVCRA UNIVERSITY LIBRARY

Literature Cited

Brandts, J.F., Halvorson, H.R., and Brennan, M. 1975. Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry* 14:4953-4963.

Creighton, T.E. 1979. Electrophoretic analysis of the unfolding of proteins by urea. *J. Mol. Biol.* 129:235-264.

Creighton, T.E. 1980. Kinetic study of protein unfolding and refolding using urea-gradient electrophoresis. *J. Mol. Biol.* 137:61-80.

Creighton, T.E. 1992. Protein Folding. W.H. Freeman, New York.

Creighton, T.E. and Shortle, D. 1994. Electrophoretic characterization of the denatured states of staphylococcal nuclease. *J. Mol. Biol.* 242:670-682.

Goldenberg, D.P. 1989. Analysis of protein conformation by gel electrophoresis. *In* Protein Structure: A Practical Approach (T.E. Creighton, ed.) pp. 225-250. IRL Press, Oxford.

Goldenberg, D.P. and Creighton, T.E. 1984. Gel electrophoresis in studies of protein conformation and folding. *Anal. Biochem.* 138:1-18.

Hollecker, M. and Creighton, T.E. 1982. Effect on protein stability of reversing the charge on amino groups. *Biochim. Biophys. Acta* 701:395-404.

Kim, P.S. and Baldwin, R.L. 1990. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59:631-660.

Klemm, J.D., Wozniak, J.A., Alber, T., and Goldenberg, D.P. 1991. Correlation between mutational destabilization of phage T4 lysozyme and increased unfolding rates. *Biochemistry* 30:589-594.

Lyubimova, T., Caglio, S., Gelfi, C., Righetti, P.G., and Rabilloud, T. 1993. Photopolymerization of polyacrylamide gels with methylene blue. *Electrophoresis* 14:40-50.

Matthews, C.R. 1993. Pathways of protein folding. *Annu. Rev. Biochem.* 62:653-683.

McLellan, T. 1982. Electrophoresis buffers for polyacrylamide gels at various pH. *Anal. Biochem.* 126:94-99.

Pace, C.N., Shirley, B.A., and Thomson, J.A. 1989. Measuring the conformational stability of a protein. *In* Protein Structure: A Practical Approach. (T. E. Creighton, ed.) pp. 311-330. IRL Press, Oxford.

Privalov, P.L. 1989. Thermodynamic problems of protein structure. *Annu. Rev. Biophys. Chem.* 18:47-69.

Schellman, J.A. 1978. Solvent denaturation. *Biopolymers* 17:1305-1322.

Schellman, J.A. 1987. The thermodynamic stability of proteins. *Annu. Rev. Biophys. Chem.* 16:115-137.

Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* 23:121-282.

Key References

Creighton, 1979. See above.

The original description of urea-gradient gels; includes extensive discussion of most of the important parameters and considerations and detailed analysis of the relationships between transition rates and electrophoretic patterns.

Goldenberg and Creighton, 1984. See above.

General review of gel electrophoresis and applications to studies of protein folding, including urea gradients.

Contributed by David P. Goldenberg
University of Utah
Salt Lake City, Utah

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Analytical Ultracentrifugation

Analytical ultracentrifugation is the grandfather of biochemical methods, with a 70-year history of development and application (Svedberg and Pederson, 1940; Schachman, 1959). Recently, the attention of protein scientists and molecular biologists has been drawn to macromolecular interactions, an area where analytical ultracentrifugation excels, resulting in a renewed interest in this technology (Harding et al., 1992; Ralston, 1993; McRorie and Voelker, 1993; Schuster and Laue, 1994). As this unit will describe, analytical ultracentrifugation is one of the most powerful, though as yet underexploited, techniques available to molecular biology and biochemistry.

Two different but complementary methods of sample analysis are possible using an analytical ultracentrifuge: sedimentation velocity and sedimentation equilibrium (Table 7.5.1). Each encompasses a range of techniques suitable for characterizing everything from crude mixtures to highly purified proteins. Sedimentation velocity provides such hydrodynamic information as size, shape, and, when combined with diffusion measurements, molecular weight. Sedimentation equilibrium provides thermodynamic information about molecular weight, stoichiometry, association energy, and nonideality.

Both sedimentation methods are understood from first principles. Thus, no standards are

Table 7.5.1 Methods of Analytical Ultracentrifugation

Method	Category	Information obtained and notes
Velocity	Direct fitting to transport equations	Modern way to obtain *s* (and the diffusion coefficient) for a single solute. When combined with molecular weight, *s* yields information concerning the size and shape of the protein (Philo, 1994).
	g(*s*)–particle size distribution analysis	Powerful method for analyzing complicated mixtures of molecules. Holds promise for analyzing associating systems (Stafford, 1992; Stafford, 1994). By increasing the gravitational field over the course of an experiment, fractionation of incredibly complicated mixtures may be achieved (Mächtle, 1988).
	Differential sedimentation	Sensitive means of detecting small (0.005%) differences in sedimentation due to conformational changes (Richards and Schachman, 1957).
	Active enzyme sedimentation	Means of characterizing an impure enzyme (Cohen et al., 1971) using a chromophoric assay and monitoring the movement of color development as the enzyme sediments.
	Van Holde and Weischet extrapolation	Graphical means of removing the effects of diffusion, thus improving resolution of a paucidisperse solution (Van Holde and Weischet, 1978).
Equilibrium	Short-column analysis (750 µm, 15 µl samples)	Useful for quick surveys of molecular weight, association properties, and nonideality (Yphantis, 1960; Laue, 1992).
	Longer-column analysis (3 mm, 110 µl samples)	Provides higher precision than short-column; useful for low-molecular-weight solutes and analysis of heterogeneity (Yphantis, 1964).
Miscellaneous	Extinction coefficient	Combines absorbance and refractive optical reading for precise measurement of extinction coeffiecient. Sensitive test for sample homogeneity.
	Tracer sedimentation	Very sensitive and selective for characterizing interactions between purified proteins.
	Diffusion coefficient	Obtained at low speed.

BCU IASI - CENTRAL UNIVERSITY LIBRARY

Copyright © 1996 by John Wiley & Sons, Inc.

required to interpret data and the same values should be obtained for parameters (e.g., molecular weight and $s_{20,w}$) from preparation to preparation and from laboratory to laboratory. Of the two methods, sedimentation velocity provides a greater degree of sample fractionation, and it is uniquely suited to particle size distribution analysis (see Anticipated Results for definitions and further discussions). The standardized sedimentation coefficient ($s^0_{20,w}$; measured in Svedberg units, S) is a constant for a given molecule; it therefore is often used as a species descriptor (e.g., 30S subunit, 5S RNA).

The reasons behind the revival of analytical ultracentrifugation reveal many of its virtues. Fundamental information concerning sample purity, native molecular weight, association energies, and molecular size and shape are all readily obtainable using this technique. The methods are not fussy with regard to solvent conditions, and can be applied to any sort of protein (fibrous or globular, large or small) over an enormous range of molecular weights (<100 to >10,000,000 g/mol). Relatively small quantities of materials are required for analysis and, because the techniques are nondestructive, samples can be recovered afterwards. Moreover, great experimental precision can be anticipated, even by a novice experimenter.

The disadvantages of sedimentation analysis are different for the velocity and equilibrium methods. Sedimentation velocity analysis requires relatively large sample volumes (450 µl) and, being a broad-zone method, exhibits poor sample fractionation when compared with various chromatographic and electrophoretic methods (Table 7.5.2). Subtle changes readily apparent upon electrophoresis may not be revealed by sedimentation. Another consequence of the technique's broad-zone nature is that only the slowest-sedimenting species can be examined in a pure state; all other boundaries will be composed of mixtures of components.

The interpretation of sedimentation equilibrium data requires that highly purified proteins be used. For the analytical ultracentrifuge, relatively high concentrations of proteins (0.1 to 1.0 mg/ml) are required for detection (Table 7.5.3). Even if these criteria are met, complicated association schemes (e.g., highly nonideal monomer:dimer:tetramer:octamer associations) are difficult to analyze in great detail. Finally, neither sedimentation velocity nor sedimentation equilibrium is suitable for kinetic analysis.

Even with these disadvantages, sedimentation analysis compares favorably with alternative methods (Table 7.5.2). There are three classes of methods that can be compared with sedimentation: (1) other primary methods (light scattering and osmotic pressure), (2) spectroscopic analysis for obtaining thermodynamic information, and (3) secondary methods built around electrophoresis and column chromatography. Both light scattering and osmotic pressure are well-grounded in thermodynamics and provide accurate estimates of molecular weight. However, neither affords any fractionation of the solution and neither provides hydrodynamic information equivalent to that from sedimentation velocity. Both are subject to problems with contamination (osmotic pressure from small particles and light scattering from large ones), though this sensitivity can be useful for detecting such contaminants. Finally, both require larger quantities and higher concentrations of sample than does sedimentation.

Spectroscopic methods—e.g., fluorescence, absorbance, circular dichroism (CD), nuclear magnetic resonance (NMR), and electron spin resonance (ESR)—are often useful for examining sample purity and for conducting binding studies. For sample purity, these methods are usually applied to detect the presence of specific contaminants (e.g., detecting protein in a DNA sample by absorbance). The concentrations needed to conduct a study range from quite low (fluorescence) to rather high (NMR). In addition to sample purity, spectroscopy often is used phenomenologically to monitor associations, including macromolecular interactions. Because of their respective sensitivities, the concentration range accessible to these methods is quite variable. All spectroscopic methods suffer from certain common problems. First, because they are phenomenological, it is not guaranteed that binding will lead to the production of an appropriate signal. Hence, it is possible for an interaction to occur that gives rise to no measurable spectroscopic change. Moreover, if the stoichiometry of binding is >1, it is possible for different binding events to give rise to proportionately different signals. In these cases, the stoichiometry of binding will be nearly impossible to determine and estimates of the strength of binding may be seriously in error.

There are several chromatographic and electrophoretic methods available for making measurements similar to those from sedimentation. These fractionate the sample and therefore are extremely useful for detecting contam-

BCU IASI - CENTRAL UNIVERSITY LIBRARY

Table 7.5.2 Alternative Analytical Methods

Class	Method	Notes
Primary methods No standards required; data interpretation grounded in physical first principles; no sample fractionation.	Scattering	Light and low-angle X-ray can yield molecular weight and radius of gyration. For mixtures, weight-average molecular weights and *z*-average radii are obtained. (Laue, 1995; Yphantis, 1964). Sensitive to large contaminants. Small cuvettes reduce sample volume, but sensitivity is limited to relatively high concentrations (>1 mg/ml). Particle scattering (e.g., neutrons) is very expensive but yields details about particular aspects of structure.
	Osmotic pressure (other colligative methods)	Useful for determining molecular weights of small molecules. For mixtures, number-average molecular weights are obtained. Sensitive to low-molecular-weight contaminants. Relatively high concentrations (>1 mg/ml) are required. Vapor pressure osmometers require small sample volumes.
Spectroscopic methods Primarily useful phenomenologically; sample size and sensitivity vary from method to method; no sample fractionation.	Absorbance	Useful for detecting particular contaminants. Simple to use. Sensitivity to contaminants depends on their extinction coefficient at the wavelength of interest. Changes in absorbance with binding can be used to monitor macromolecular interactions with varying success.
	Fluorescence	Very sensitive. Both luminescence and quenching of fluorescence useful for detecting contaminants. Changes in fluorescence can be used to monitor macromolecular interactions with varying success.
	Circular dichroism	Not generally as useful as other spectroscopic methods for detecting contaminants. Can be used to monitor macromolecular interactions with varying success. Instrumentation expensive.
	Nuclear magnetic resonance, electron spin resonance	Not generally as useful as other spectroscopic methods for detecting contaminants. Can be used to monitor macromolecular interactions with varying success. Requires high concentrations of material and instrumentation expensive.
Secondary methods Require standards for interpretation; can be particularly useful for detecting contaminants; most are used as "thin-zone" methods, which limits their usefulness in characterizing interactions.	Column chromatography	Depending on column matrix, can be sensitive to size, charge, or specific binding. In studying binding, useful qualitatively but less so for quantitative work. Fractionation can be excellent. Ability to measure contaminants depends on detector used to monitor effluent, but can be exquisitely sensitive. A modified detector in which the matrix is attached to the detector is at the heart of the Pharmacia Biotech Biacore and related devices.
	Electrophoresis	Several methods based on electrophoretic mobility exist (e.g., gel electrophoresis, capillary zonal electrophoresis, and isoelectric focusing). Fractionation is usually superb; sensitivity depends on detector. Because of the high degree of sample fractionation, only limited, qualitative information about macromolecular interactions is obtainable.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 7.5.3 Methods of Solute Detection

Method of detection	Use[a]	P[b]	A[c]	S[d]	D[e]	Notes
Absorbance	A, P	G	F	G	F	Sensitivity depends on solute's extinction coefficient. Discrimination depends on separation between absorbances.
Refractive	A	E	E	G	P	Accuracy and precision cannot be matched by other techniques. Method of choice for nonabsorbing solutes or buffers that absorb in the UV.
Fluorescence	P	F	F	E	E	Specific labels can provide excellent sensitivity and superb discrimination, but the linearity of the concentration determination is subject to several types of error.
Enzyme assay	P, A	F	F	E	E	Excellent discrimination and sensitivity. Can be used to determine which form of an enzyme is active. Precision and accuracy can be quite good or rather poor, depending on the assay. A band sedimentation method for the analytical ultracentrifuge works quite well, but requires that a suitable colorimetric assay be available.
Radiolabel	P	G	F	E	E	Superb discrimination and sensitivity can be attained. Accuracy of the concentration determination can range from very good to very poor depending on the particulars of the radiolabeling.
Other	P	—	—	—	—	Any assay (RIA, ELISA, bioassay) that yields a signal proportional to the concentration may be used for determining molecular weight and $s_{20,w}$. Of particular interest are methods in which the sample is fractionated (e.g., by gels or HPLC) subsequent to sedimentation, which allows detection of interactions between components in extremely complicated mixtures. Typically, the accuracy of these methods is not sufficiently high to allow further analysis (i.e., determination of association constants).

[a]Designates the instrument that may be used: A, analytical ultracentrifuge; P, preparative centrifuge with subsequent microfractionation. Where both types of instruments may be used, the better of the two is listed first. In general, the analytical ultracentrifuge offers improved accuracy in the determination of radial position and concentration, which is necessary for determining accurate association constants and stoichiometries. The preparative centrifuge offers lower cost and, in general, greater solute sensitivity and discrimination, but lower precision and accuracy.

[b]Precision of detection. Usually given as a standard error of replicates, the precision of a detection scheme may be very high, yet the accuracy may be poor if systematic error is present. E, excellent; G, good; F, fair; P, poor.

[c]Accuracy of detection. This depends on the lack of systematic errors (e.g., nonlinearity of the response of the detector with concentration).

[d]Sensitivity to solute.

[e]Discrimination between solutes.

inants. Most of these methods are now highly refined, are easy to use, and can examine a wide range of sensitivities. However, with regard to sample characterization, they are less rigorous than sedimentation. This stems from the fact that all are secondary methods, requiring standards for calibration and comparison. For example, gel permeation chromatography is often used to estimate molecular weights, radii, and stoichiometries. However, if a protein does not conform to the properties of the standards (e.g., if the unknown is asymmetric and the standards are all spherical), erroneous conclusions may be reached with this approach. A second, more fundamental, problem arises when these methods are used to examine macromolecular interactions. Both chromatography and electrophoresis usually entail the monitoring of thin zones of material. Because the concentration in the zone varies both spatially and with time, it is impossible to relate the zone shape or position to any association parameters in a rigorous fashion.

PLANNING AN EXPERIMENT

The principle measurement obtained in any sedimentation experiment is the concentration as a function of radial position. Knowledge of the concentration distribution, along with the rotor speed and temperature, is needed to obtain a molecular weight or an $s_{20,w}$. There are several means of determining the concentration distribution (Table 7.5.3). Analyses may be conducted using either a preparative or an analyt-

BCU IAS/CENTRAL FIRST LIBRARY

ical ultracentrifuge. It is worth taking a moment to consider which instrument will be most useful for elucidating a particular problem. To do this, it is necessary to have a clear notion of what information is being sought from the centrifugation experiment. If all that is needed is a rough estimate of the molecular weight to confirm a stoichiometry, then a preparative centrifuge will be sufficient. On the other hand, if association constants or accurate sedimentation coefficients are needed, an analytical ultracentrifuge will be required.

The biggest difference between an analytical and a preparative ultracentrifuge is the analytical instrument's ability to monitor the concentration distribution as a sample is spinning. To do this, a special cell and rotor are used that permit light to pass through the sample. Light sources and detectors are arranged to monitor the cell contents throughout an experiment. Data acquisition is automated, and a large number (500 to 2000) of data points are available across the cell (1.5 cm), yielding the high radial resolution needed for the analysis of interacting systems. At present, absorbance and refractive detection are available for the XLA analytical ultracentrifuge (Beckman), and prototype fluorescence optics are under development. Data acquisition requires 5 to 30 sec per cell, and up to four cells may be used.

The ability to acquire data rapidly and with high radial resolution has led to improvements in both sedimentation equilibrium and velocity analysis. For sedimentation equilibrium, there has been a resurgence of interest in short-column methods (Yphantis, 1960; Laue, 1992), whose small sample volumes (15 µl), rapid equilibrium (typically ~1 hr), and large throughput (up to 16 samples simultaneously) are well-suited to addressing biochemical problems. The development of time-derivative sedimentation velocity has improved the sensitivity of the refractive optical systems nearly 100-fold, and promises to revolutionize the range of systems suitable for sedimentation analysis. Experimental protocols in which the rotor speed is accelerated over the course of an experiment (i.e., gravitational sweep sedimentation) have been described (Stafford, 1992; Machtle, 1988) for analyzing and characterizing polydisperse samples. The fractionation range and accuracy of this method far exceed those of any other technique presently available.

A good preparative ultracentrifuge (either a floor or tabletop model) may also be used for analytical ultracentrifugation, as confirmed by the long-time success of sucrose gradient analysis. In many cases, these methods offer great advantages because the sample is analyzed after centrifugation by taking fractions (i.e., radial slices) from the centrifuge tube. Because of this, any detection scheme may be used to determine the concentration distribution. This means that unparalleled sensitivity and selectivity are afforded by schemes that use the preparative ultracentrifuge. Moreover, each fraction may be further fractionated—e.g., by high-performance liquid chromatography (HPLC) or gel electrophoresis—yielding good analytical sedimentation data for complicated mixtures. Detailed protocols for sedimentation velocity and sedimentation equilibrium in a preparative ultracentrifuge are available, and a superb microfractionator is available (from Brandel) specifically for conducting sedimentation analyses (Attri and Minton, 1986).

However, it must be recognized that the information obtained from a preparative centrifuge is a snapshot of the cell at the time of deacceleration (assuming no mixing occurs during deacceleration). This precludes the use of the elegant time-derivative methods that have been developed for the analytical machine. Moreover, the time delay between stopping the run and sampling the cell allows diffusion to distort the concentration gradient, leading to inaccuracy in its determination. This problem becomes severe when steep gradients are involved. In general, the radial resolution for the preparative techniques is not as high as with the analytical ultracentrifuge.

The concentration of protein needed for an experiment depends on what question is being addressed. If one merely seeks the molecular weight or sedimentation coefficient of a solute, then any convenient concentration may be used. However, if one is interested in knowing the association state of a sample under particular conditions (e.g., at concentrations found in vivo, or in an assay), the choice of detection scheme may become critical. In this case, one must work out what concentration is needed and what detection schemes are suitable for the concentration range, and match this to an entry in Table 7.5.3. It is essential to realize that any experimental protocol must examine the sedimentation behavior over a range of concentrations to take account of thermodynamic nonideality (sedimentation equilibrium) and hydrodynamic effects (sedimentation velocity; Laue et al., 1992).

The stability and accuracy of the temperature is a critical experimental parameter. The

BCM LIBRARY

XLA ultracentrifuge can be operated at temperatures between 0° and 40°C. Although this is convenient for most biochemical work, there are some studies (e.g., thermal denaturation) where a wider range is desired. In this case, one is limited to using an appropriately configured preparative centrifuge. Regardless of the temperature used, temperature stability is important in preventing convection. The accuracy of the temperature measurement is also important, particularly for sedimentation velocity analysis, where adjustment of a measured sedimentation coefficient (s) to standard conditions (20°C and water; the $s_{20,w}$) must be made. Methods for making these adjustments are well established and have been automated (Laue et al., 1992).

The rotor speed should be chosen carefully, as the gravitational field (g) increases as its square ($g = \omega^2 r$, where $\omega = \pi \times \text{rpm}/30$ and r is the rotor radius). Different criteria are used to choose the rotor speed for sedimentation velocity and for sedimentation equilibrium analyses. For a sedimentation velocity experiment, two options can be considered: (1) gravitational sweep, in which the rotor speed is increased periodically throughout the experiment, and (2) s, where a single rotor speed is chosen for the analysis. Gravitational sweep is particularly useful for the initial characterization of a system where it is useful to get a general idea of how broad the particle size distribution is. In this method, the rotor is accelerated to a low speed (e.g., 3000 rpm) and several (six to ten) concentration distributions are collected to determine if there are any fast-moving boundaries. Next, the rotor speed is accelerated twofold, and data collection again undertaken. This process is continued until the maximum rotor speed (60,000 rpm) is attained, following which data are acquired until all boundaries have sedimented at least three-quarters of the way down the cell. Using this scheme, a sample can be "scanned" for the presence of an extraordinarily broad range of particle sizes (from greater than 1000S to ~1S).

The second method of analysis acquires data at a fixed rotor speed. This is useful for samples that have one or only a few types of proteins with similar sizes (within an order of magnitude). If one is trying to resolve two proteins of similar size, the highest rotor speed that allows a sufficient quantity of data to be obtained should be used (i.e., such that the fastest boundary takes ~30 min to travel the length of the cell). The time-derivative data analysis routines work best with large numbers of data sets closely spaced in time, so data should be acquired frequently throughout a run.

Multiple rotor speeds should be used for sedimentation equilibrium analysis. Choosing appropriate rotor speeds is relatively simple if the molecular weight of the largest species is known ($\text{rpm} \approx 2 \times 10^6 \sqrt{M^{-1}}$), and a practical scheme for choosing an initial rotor speed is available for other cases. Low speeds are useful for detecting aggregates and high speeds are useful for detecting proteolytic fragments. In general, rotor speeds should be used that keep the reduced molecular weight, σ, between 2 and 12. An estimate of σ can be calculated from the molecular weight:

$$\sigma = M(1 - \bar{v}\rho)\,\omega^2/RT$$

where $\bar{v}$ is the partial specific volume (~0.72 ml/g is a good rough estimate), ρ is the solvent density (1 g/ml for the dilute aqueous buffers usually used in biochemistry), ω is the rotor angular velocity (see above), R is the gas constant (8.31×10^7 erg/mol°), and T is the absolute temperature. Even a crude estimate of σ will suffice. The range of rotor speeds used in an analysis should vary σ over a 4-fold or greater range, if possible. The experimental protocol should involve incrementing rather than decrementing the rotor speed, as (because of diffusion) it takes much longer for a rotor to reach equilibrium when approaching it from a higher speed. There are many other practical tips for choosing the rotor speed, estimating σ, and designing an appropriate experimental protocol (Ralston, 1993; Laue, 1992).

In general, conducting an experiment with an analytical ultracentrifuge is very simple. The usual cause of difficulty is the sample, either in terms of its purity (for experiments trying to extract thermodynamic details) or its behavior (for previously uncharacterized biochemical systems). Problems often encountered for either sedimentation velocity or sedimentation equilibrium, along with suggestions for resolving them, are discussed in Table 7.5.4.

ANTICIPATED RESULTS

Different, yet complementary, information is obtained from sedimentation velocity and sedimentation equilibrium experiments. For sedimentation velocity, the rate of movement of the boundary yields the sedimentation coefficient, s. Straightforward adjustment of this measured value provides $s_{20,w}$, the expected sedimentation coefficient for the molecule if it were sedimented in pure water at 20°C (Laue et al., 1992). A second adjustment is made by

BCU IAS CENTRAL UNIVERSITY LIBRARY

Table 7.5.4 Troubleshooting Guide for Analytical Ultracentrifugation

Problem	Possible cause	Solution
Velocity method		
No boundary or very broad boundary	Convection	May be caused by thermal instability or mechanical instability (rotor vibration)
	Sample heterogeneity (aggregates or fragments)	Clean sample by gel filtration
	Loss of sample to cell walls	Use centerpiece made of different material.
	Interacting system (concentration-dependent; may show pressure dependence)	Analyze by sedimentation equilibrium
Too few data points; boundary reaches base in <30 min	Rotor speed too high	Shake cell and rerun at lower rotor speed
Boundary that moves too slowly	Rotor speed too low	Shake cell and rerun at higher rotor speed
	Presence of proteolytic fragments	Fractionate sample by gel filtration
	Improper sample dialysis	Buffer component not equilibrated in sample and reference channels
Equilibrium method		
Decrease in molecular weight with increased concentration	Nonideality; little dependence of rotor speed on molecular weight	Use a nonideal model in nonlinear least squares analysis; extrapolate to zero protein concentration. Information about molecular size and charge available.
	Sample heterogeneity; molecular weight also decreases with increased rotor speed	Fractionate sample
Increase in molecular weight with increased concentration, but little change as rotor speed is varied	Mass-action association	Characterize by nonlinear least-squares analysis
Somewhat higher molecular weight than expected; slow loss of material	Slow aggregation; flocculant formation (data are often noisier than usual; inconsistent molecular weights are obtained from one sample to the next)	Use sedimentation velocity analysis to detect this as a fast boundary or sloping plateau concentration. Only cure is to find buffer conditions that stabilize protein solubility. For intracellular proteins, be sure a reducing agent (2-ME or DTT) is present.
Somewhat lower molecular weight than expected	Proteolysis (e.g., molecular weight drops with time)	Use protease inhibitors in sample; refractionate by gel filtration.
	Improper dialysis of sample prior to analysis	Be sure sample and reference buffers are at dialysis equilibrium prior to run. Use centrifugal gel filtration for rapid equilibration.

measuring the sedimentation coefficient of the boundary at several different protein concentrations, then extrapolating the values of $s_{20,w}$ to zero protein concentration to get $s^0_{20,w}$. For a nonassociating protein, $s^0_{20,w}$ decreases with increasing concentration; this decrease occurs gradually if the protein is spherical (by ~1% per mg/ml) and more strongly if it is asymmetric. In addition to gaining insight into the shape of a protein, there are three reasons for adjusting to standard conditions. First, $s^0_{20,w}$ is a practical, descriptive constant for a protein that can be used to compare preparations made under different conditions or in different laboratories. Second, determination of $s^0_{20,w}$ provides a first-principle means of detecting changes in the hydrodynamics of the protein (see below). Third, an increase in $s^0_{20,w}$ with increasing concentration is a useful diagnostic for mass action association of a protein.

Sedimentation velocity is also useful for characterizing more complicated mixtures of molecules (Stafford, 1992). In these cases sedimentation coefficient profiles are obtained. For samples containing only a few boundaries (paucidisperse samples), appropriate extrapolation of the data to infinite time provides higher precision and improved resolution of the

BRANDEIS CENTRAL UNIVERSITY LIBRARY

boundaries (Van Holde and Weischet, 1978; Stafford, 1994).

In addition to $s^0_{20,w}$, the Stokes radius, R_S, of a particle may be obtained from sedimentation analysis (Laue et al., 1992). Models can be applied in which R_S is analyzed in terms of either the extent of hydration or the degree of asymmetry (Laue et al., 1992). Such analyses often provide the first insight into the shape of a protein.

Equilibrium sedimentation provides thermodynamic information about a sample. Analyzed in the simplest way, the data can be used to obtain the molecular weight. For a mixture of molecules, various average molecular weights (number, weight, and z) may be calculated (Yphantis, 1964; Laue, 1995). By measuring the molecular weight in a denaturing solvent (e.g., 6 M guanidine·HCl; *APPENDIX 3A*) and in a physiological solvent, the stoichiometry of the native molecule can readily be determined (Laue, 1992). More detailed analysis of the concentration distributions can provide energies of assembly and nonideality coefficients (McRorie and Voelker, 1993).

TIME CONSIDERATIONS

For a typical run with the XLA ultracentrifuge involving three cells per rotor, sample preparation and cell loading requires ~1 hr. For sedimentation velocity analysis, a run requires ~2 hr (once the rotor is at the desired temperature). Gravitational sweep analysis requires somewhat longer. The time required to conduct a sedimentation equilibrium analysis is discussed below. In any case, operation of the centrifuge, including data acquisition, is completely automated. After the run, sample recovery and cell cleaning require ~1 hr. For maximum efficiency, it is worthwhile to have two complete sets of cells so that one may be cleaned and loaded while the other is being run. Data analysis for sedimentation velocity experiments is computerized and rather straightforward, and usually requires <1 hr.

The time required to conduct a sedimentation equilibrium experiment depends a great deal on the particulars of the experimental setup. The number of samples that can be analyzed simultaneously depends on the choice of centerpieces, so that up to sixteen samples may be analyzed simultaneously. Cell loading still only requires ~1 hr. The time needed to reach sedimentation equilibrium depends on the solvent viscosity, the diffusion coefficient of the protein, and the distance between the meniscus and the base of the solution column. In general, the time increases linearly with the viscosity, as the cube root of the molecular weight, and as the square of the column height of the solution. Shorter solution columns can be accommodated in any of the centerpieces. For a typical aqueous buffer and a 50,000-molecular-weight protein, the time needed to reach equilibrium is ~9 hr for a 3-mm solution column (~110 µl of sample) and ~1 hr using a short column centerpiece (~15 µl of sample). Because a typical experimental protocol requires equilibrium at three or four rotor speeds, a full 1 or 2 days are needed to complete data acquisition.

Computer programs have automated equilibrium data analysis. Even so, the time needed to analyze equilibrium data depends on the level of detail of the scrutiny. If only an average molecular weight is required, this can be accomplished in a few minutes. On the other hand, determination of both the stoichiometry and energetics of a self-associating protein may require several days of analysis if no previous characterization has been done. Both the level of detail required to answer a biological question and the experience of the researcher will affect the time needed to complete an analysis.

For either velocity or equilibrium sedimentation, there are no time-critical steps once the centrifuge run has started. Zealots often begin data analysis during the centrifuge run. Although this is not strictly necessary, it is particularly useful when examining a previously uncharacterized system as it allows adjustment of the experimental conditions to maximize the quality of the data.

Methods using a preparative centrifuge can handle a larger number of samples in a run (depending on which rotor is used). Likewise, the time required for preparation will vary, though it will be on the same order of magnitude as for the XLA. Unlike with the XLA, data acquisition is not fully automated in that the radius of each fraction must be calculated on the basis of its volume and the sample-tube geometry (Attri and Minton, 1986). Likewise, assaying the fractions may or may not require considerable time.

HELP FOR THE NOVICE

It is widely recognized that undertaking sedimentation analysis for the first time can be a daunting task. Fortunately, the most difficult parts—including data acquisition, data analysis, and data interpretation—have been automated. Commercial computer programs for these tasks are available from Beckman Instru-

RCU IASI CENTRAL UNIVERSITY LIBRARY

ments. In addition, there is an Internet bulletin board (rasmb@bbri.eri.harvard.edu) where analytical sedimentation is discussed and a wide variety of computer programs are available free of charge. Often, though, it is not the process of conducting analytical ultracentrifugation that is a problem for biochemists, but rather their uncertainty about the solution physical chemistry of proteins. Fortunately, the cadre of researchers working with the ultracentrifuge welcome the opportunity to help newcomers to the field. For further information contact the author by e-mail at tml@hopper.unh.edu.

LITERATURE CITED

Attri, A.K. and Minton, A.P. 1986. Technique and apparatus for automated fractionation of the contents of small centrifuge tubes: Application to analytical ultracentrifugation. *Anal. Biochem.* 152:319-328.

Cohen, R. and Mire, M. 1971. Analytical-band centrifugation of an active enzyme-substrate complex. 1. Principle and practice of the centrifugation. *Eur. J. Biochem.* 23:267-275.

Harding, S.E., Rowe, A.J., and Horton, J.C. 1992. Analytical Ultracentrifugation in Biochemistry and Polymer Science. Royal Society of Chemistry, Cambridge.

Laue, T.M., Shah, B., Ridgeway, T.M., and Pelletier, S.L. 1992. Computer-aided interpretation of sedimentation data for proteins. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science (S.E. Harding, A.J. Rowe, and J.C. Horton, eds.) pp. 90-125. Royal Society of Chemistry, Cambridge.

Laue, T.M. 1992. Short Column Sedimentation Equilibrium Analysis for Rapid Characterization of Macromolecules in Solution. Technical Information DS-835. Beckman Instruments, Palo Alto, Calif.

Laue, T.M. 1995. Sedimentation equilibrium as a thermodynamic tool. *Methods Enzymol.* 259:427-452.

Mächtle, W. 1988. Coupling particle size technique. A new ultracentrifuge technique for determination of the particle size distribution of extremely broad distributed dispersions. *Die Angew. Makromol. Chem.* 162:35-52.

McRorie, D.K. and Voelker, P.J. 1993. Self Associating Systems in the Analytical Ultracentrifuge. Beckman Instruments, Fullerton, Calif.

Philo, J.S. 1994. Measuring sedimentation, diffusion and molecular weights of small molecules by direct fitting of sedimentation velocity concentration profiles. *In* Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems (T.M. Schuster and T.M. Laue, eds.) pp. 156-170. Birkhauser, Boston.

Ralston, G. 1993. Introduction to Analytical Ultracentrifugation. Beckman Instruments, Fullerton, Calif.

Richards, E.G. and Schachman, H.K. 1957. A differential ultracentrifuge technique for measuring small changes in sedimentation coefficient. *J. Am. Chem.* 79:5324-5325.

Schachman, H.K. 1959. Ultracentrifugation in Biochemistry. Academic Press, New York.

Schuster, T.M. and Laue, T.M. 1994. Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems. Birkhauser, Boston.

Stafford, W.F. III. 1992. Boundary analysis in sedimentation transport experiments: A procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal. Biochem.* 161:70-79.

Stafford, W.F. III. 1994. Sedimentation boundary analysis of interacting systems. *In* Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems (T.M. Schuster and T.M. Laue, eds.) pp. 119-137. Birkhauser, Boston.

Svedberg, T. and Pederson, K.O. 1940. The Ultracentrifuge. Clarendon Press, Oxford.

Van Holde, K.E. and Weischet, W.O. 1978. Boundary analysis of sedimentation-velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers* 17:1387-1403.

Yphantis, D.A. 1960. Rapid determination of molecular weights of peptides and proteins. *Ann. N.Y. Acad. Sci.* 88:586-601.

Yphantis, D.A. 1964. Equilibrium ultracentrifugation of dilute solutions. *Biochemistry* 3:297-317.

KEY REFERENCES

Harding et al., 1992. See above.

Excellent book containing descriptions of many of the modern methods of analytical ultracentrifugation.

Schachman, 1959. See above.

Classic work describing many useful methods still used today.

Schuster and Laue, 1994. See above.

Up-to-date collection of methods for analytical ultracentrifugation.

Svedberg and Pederson, 1940. See above.

A classic work that is still relevant in that it provides some of the clearest descriptions available of analytical centrifugation.

INTERNET RESOURCES

www.bbri.harvard.edu/rasmb/rasmb.html

Web site for most recent programs and discussion group on analytical ultracentrifugation.

Contributed by Thomas M. Laue
University of New Hampshire
Durham, New Hampshire

B. C. U. „M. EMINESCU" IAŞI
CHITIE

Determining the CD Spectrum of a Protein

There are two requirements for a molecule or group of atoms in a molecule to exhibit a circular dichroism (CD) spectrum. The first is the presence of a **chromophore**—i.e., a group that can absorb radiation by virtue of the electronic configuration of its resting or ground state at room temperature. The energy absorbed results in a transition to a higher-energy or excited state, which has a different distribution of electrons around the nucleus. It can therefore interact with its environment in a way that differs from the ground state. In proteins, tryptophan, tyrosine, and phenylalanine are the main chromophores in the near-UV (240- to 320-nm) region; the peptide bond is the main chromophore in the far-UV (180- to 240-nm) region. Disulfide bonds and histidine residues are two other chromophores whose contribution to CD are, in general, less marked. Most chromophores exhibit more than one transition (e.g., see Fig. 7.6.1, Fig. 7.6.2, and Table 7.6.2) and their spectra are a composite of several absorption bands. It is an important feature of CD that, unlike optical rotation, the wavelength region within which spectra are observed is limited strictly to the wavelength region of the individual absorption bands. This confers greater specificity on spectra and allows a certain degree of assignment to particular chromophores.

The second requirement for CD is that the chromophore be in, or closely associated with, an **optically asymmetric environment**. The chromophores in proteins are themselves not chiral and exhibit no optical activity. The phenolic group of tyrosine, for example, exhibits a CD spectrum only because it is connected to an optically asymmetric carbon atom. However, when the same group is packed in a folded protein in an environment that is asymmetric with respect to polarity, or is interacting through the phenolic hydroxyl, it exhibits a different CD spectrum that is specific to the particular environmental influences acting on the chromophore. In practice, the latter spectrum is usually much more intense than that for free tyrosine and may also differ in shape. When the peptide-bond chromophore is part of a regularly folded structure, with particular backbone angles and hydrogen-bond interactions as in an α-helix or β-sheet, it becomes closely associated with a conformationally asymmetric structure—i.e., one that cannot be superimposed on its mirror image. A CD spectrum for the protein will be generated that is a composite of the individual spectra corresponding to each of the peptide-bond absorption transitions (Fig. 7.6.1). The intensity of the spectrum may be further affected by interaction between neighboring chromophores (Bayley, 1980).

Plane-polarized radiation comprises two circularly polarized vectors of equal intensity, one right-handed and the other left-handed (Fig. 7.6.3A), which are separately measured in the CD spectrometer by means of a photoelastic modulator. A chromophore situated in an optically symmetrical environment will normally absorb the two components equally so that, when recombined after passing through a solution of the chromophore, they result once again in radiation oscillating in a single plane. A chromophore situated in an optically asymmetric environment, however, will absorb each of the two components to a different extent, the difference being ΔA. When recombined, the resultant vector describes an ellipse (Fig. 7.6.3B), the ratio of whose major and minor axes determines the **ellipticity.** The value of ΔA, and hence that of ellipticity, can be positive or negative depending on the nature of the asymmetric environment. This is exemplified by the contribution of the tyrosines in IL-1β (Fig. 7.6.4A) and of the tyrosines and tryptophans in GM-CSF (Fig. 7.6.5).

It is a general consequence of the above principles that CD spectra of molecules in solution are located in the same wavelength region as their absorption bands. For proteins this means the far-UV and near-UV regions, as well as regions extending into the visible and near infrared. These regions have their origin in and provide information about the

Copyright © 1996 by John Wiley & Sons, Inc.

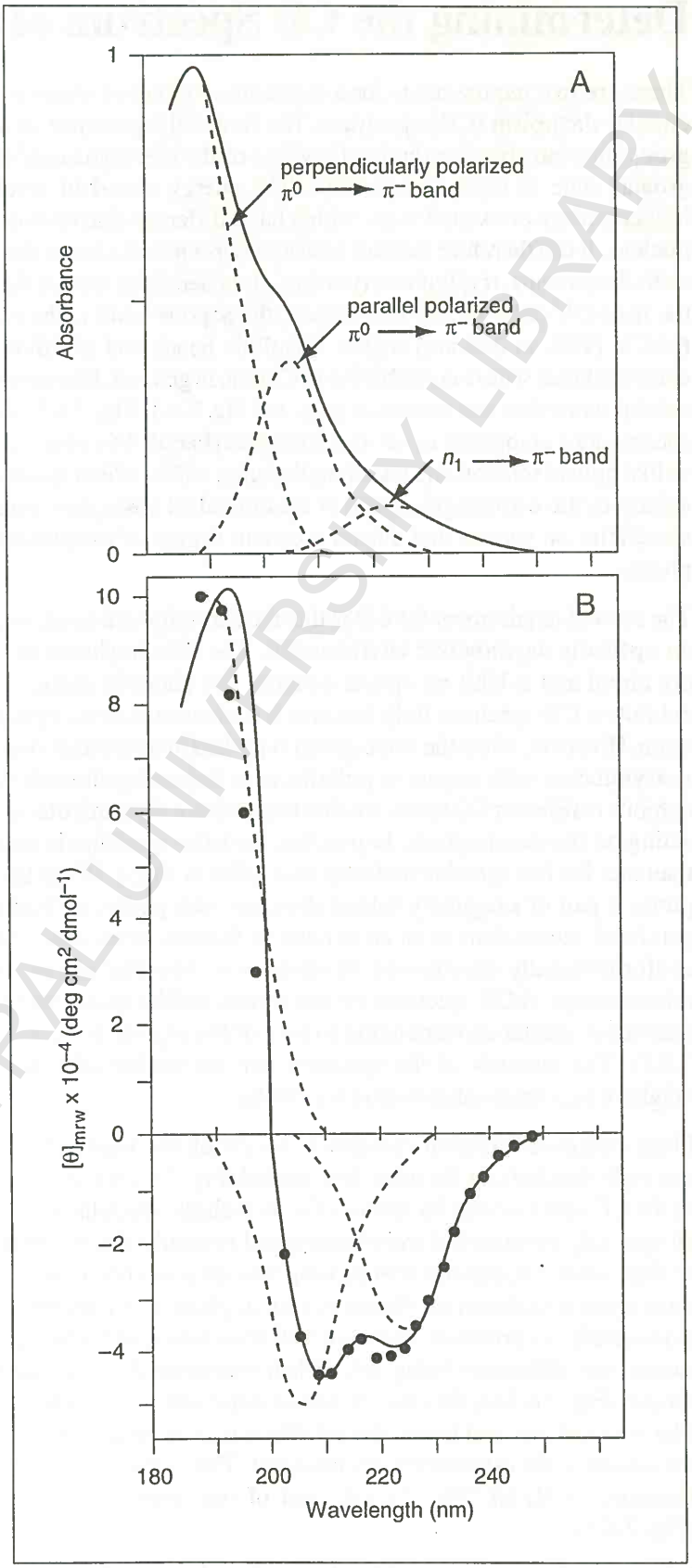


Figure 7.6.1 Relation between absorbance and circular dichroism for a polypeptide chain in the helical conformation. **(A)** The absorbance spectrum of poly-γ-methyl-L-glutamate in the α-helical form, with an assignment of its three constituent transitions. **(B)** The far-UV CD spectrum of the same compound, deconvoluted on the assumption that the shapes of the absorbance and dichroic bands are directly related (broken lines). The solid line is the observed spectrum and the filled circles the sum of the three bands. In general, the sign of a CD band may be either positive or negative, and its intensity does not necessarily follow that of the absorbance band. (From Holzwarth and Doty, 1965.)

PCU IASI / CENTRAL UNIVERSITY LIBRARY

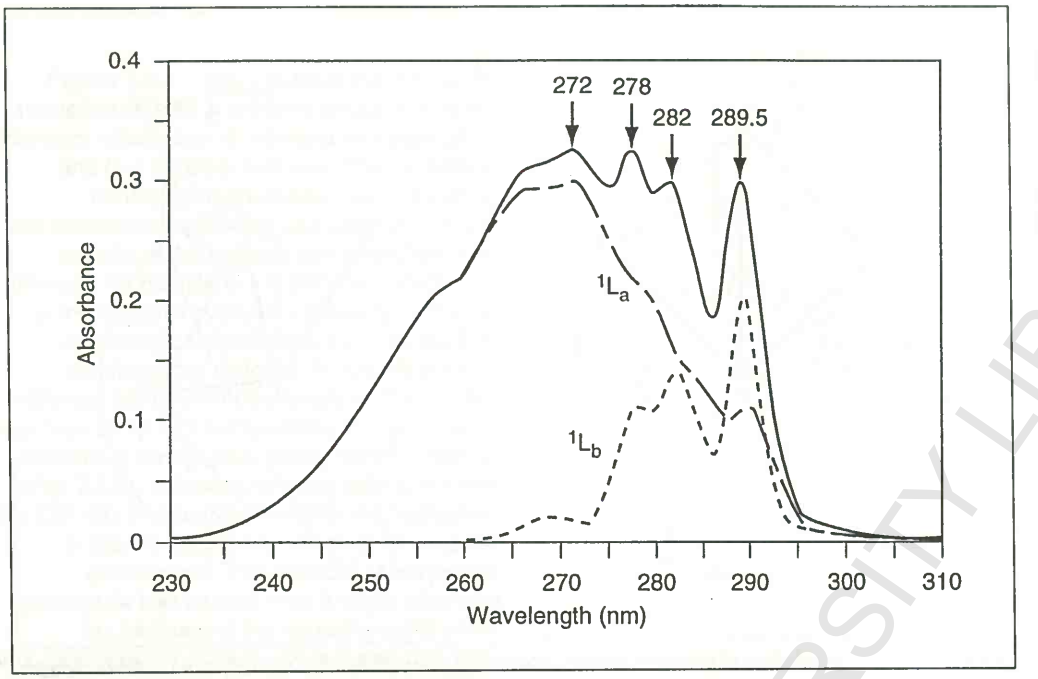


Figure 7.6.2 The two near-UV absorbance transitions for tryptophan. The near-UV absorption spectrum for *N*-stearyl tryptophan *n*-hexyl ester dissolved in methylcyclohexane at 24°C (solid line) has been resolved into the ^1L_a and ^1L_b bands. (Drawn from Strickland, 1974.)

polypeptide backbone and its conformation (far-UV), the aromatic amino acid residues and their environments (near-UV), and bound ligands such as heme and cofactors (visible and near infrared). Experimentally, the near-UV and the visible and near-infrared regions can be treated together and the protocols in this unit will therefore refer explicitly only to the collection of far- and near-UV spectra.

In both the far- and near-UV regions, CD spectra can be used empirically as "fingerprints" of a particular protein, with the spectrum resulting from the aromatic residues being rather more specific and hence diagnostic. The far-UV spectra, however, can provide information about the protein conformation in terms of its secondary structure. As for fluorescence spectroscopy or any spectroscopic method, the sample needs to be chemically pure and homogeneous.

The output of CD spectrometers is in two alternative but related types of unit. The difference in absorbance with respect to left- and right-handed circularly polarized radiation ($A_L - A_R = \Delta A$) is measured in absorbance units; ellipticity (θ) is measured in millidegrees (mdeg). These units are related by the expression $\theta = 33{,}000 \times \Delta A$ (see Fig. 7.6.3A).

Ellipticity is usually used for far-UV measurements. Because essentially every residue in a protein is associated with a peptide bond, spectra in this region can be directly compared for different proteins by calculating the mean residue ellipticity,

$$[\theta]_{mrw} = \frac{\theta \times M_{mrw}}{10 \times c \times 1}$$

where c is the protein concentration in mg/ml, l is the cell path length in cm, and M_{mrw} is the mean residue molecular weight—i.e., protein formula weight/number of residues. The dimensions of $[\theta]_{mrw}$ are deg cm^2/dmol, and the quantity is independent of molecular weight. M_{mrw} is in the range 110 to 115 for most proteins.

BOJIAN / CENTRAL UNIVERSITY LIBRARY

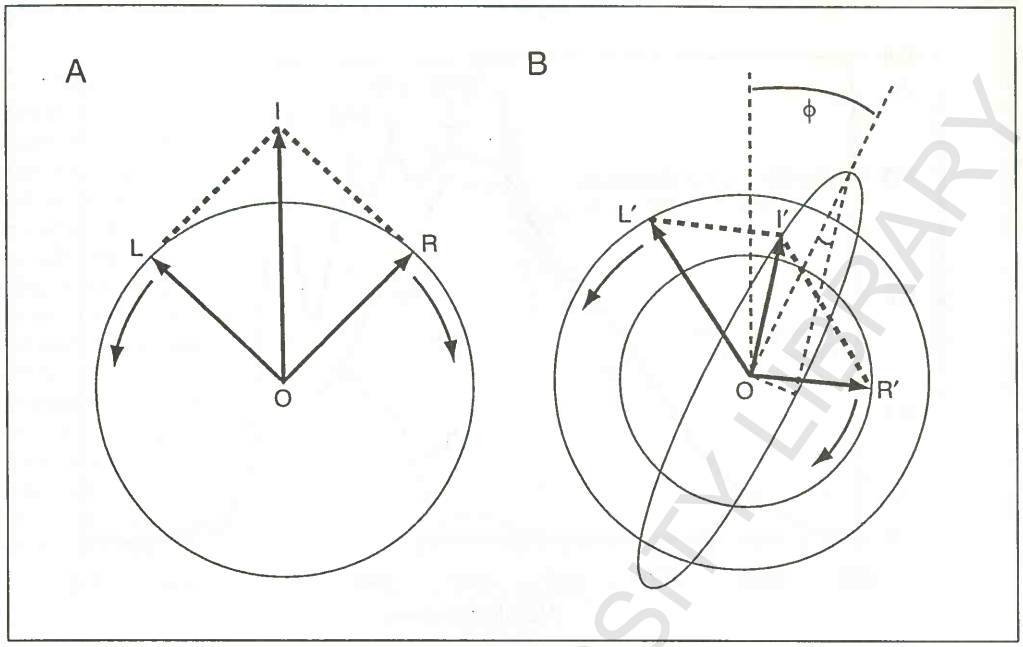


Figure 7.6.3 The relation of ellipticity to the differential absorption of circularly polarized radiation. The oscillating radiation sine wave, **OI**, is proceeding out of the plane of the paper towards the viewer. **(A)** Plane-polarized radiation is made up of left- and right-handed circularly polarized components, **OL** and **OR**, respectively. Absorption by a chromophore in a nonchiral environment results in an equal reduction in intensity of each component, whose resultant is a vector oscillating only in the vertical plane—i.e., plane-polarized radiation. **(B)** Interaction of the radiation with a chiral chromophore leads to unequal absorption, so that combination of the emerging vectors, **OL'** and **OR'**, leads to a resultant that describes an elliptical path as it progresses out of the plane of the paper. The ratio of the major and minor axes of the ellipse is expressed by tan θ, thus defining ellipticity. The major axis of the ellipse makes an angle (ϕ) with the original plane, which defines the optical rotation. This figure thus demonstrates the close relation between optical rotation and circular dichroism.

For example, for serum albumin (see Fig. 7.6.6), the ellipticity measured at 209 nm using a 1-mm path-length cell for a solution containing 0.361 mg/ml is –61.7 mdeg. Taking a mean residue molecular weight of 110, $[\theta]_{mrw} = -18.8 \times 10^3$ deg cm^2/dmol.

For small molecules, such as aromatic amino acids and model compounds, the molar difference absorption coefficient is frequently used. By analogy with absorbance spectroscopy, this quantity is given by $\Delta\varepsilon = \Delta A/(c \times l)$, where c is the concentration in mol/liter, l the path length in cm and $\Delta\varepsilon$ has the units liters/mol/cm (or cm^2/mmol). However, in the case of proteins, near-UV CD spectra are made up of contributions from tryptophan, tyrosine, and phenylalanine residues—the content of all of which vary from protein to protein. It is therefore confusing—and achieves no useful end—to present results in terms of mean residue ellipticity. Instead, dichroism should be presented as values of ΔA. For a given protein, different samples may be compared by calculating $\Delta A/(c \times l)$, with c in mg/ml.

Owing to the variety of units found in the literature and in software—e.g., mdeg and deg for ellipticity, and mm, cm or dm for path length—it is necessary to take care that experimental parameters be entered in the units requested by the machine software.

The Strategic Planning section deals with considerations regarding instrumentation and reagents for CD spectrometry. The Basic Protocol outlines the steps in recording a CD spectrum. The two support protocols explain the interpretation of CD spectra—Support Protocol 1 deals with near-UV spectra and Support Protocol 2 with far-UV spectra.

PCU IASRI CENTRAL UNIVERSITY LIBRARY

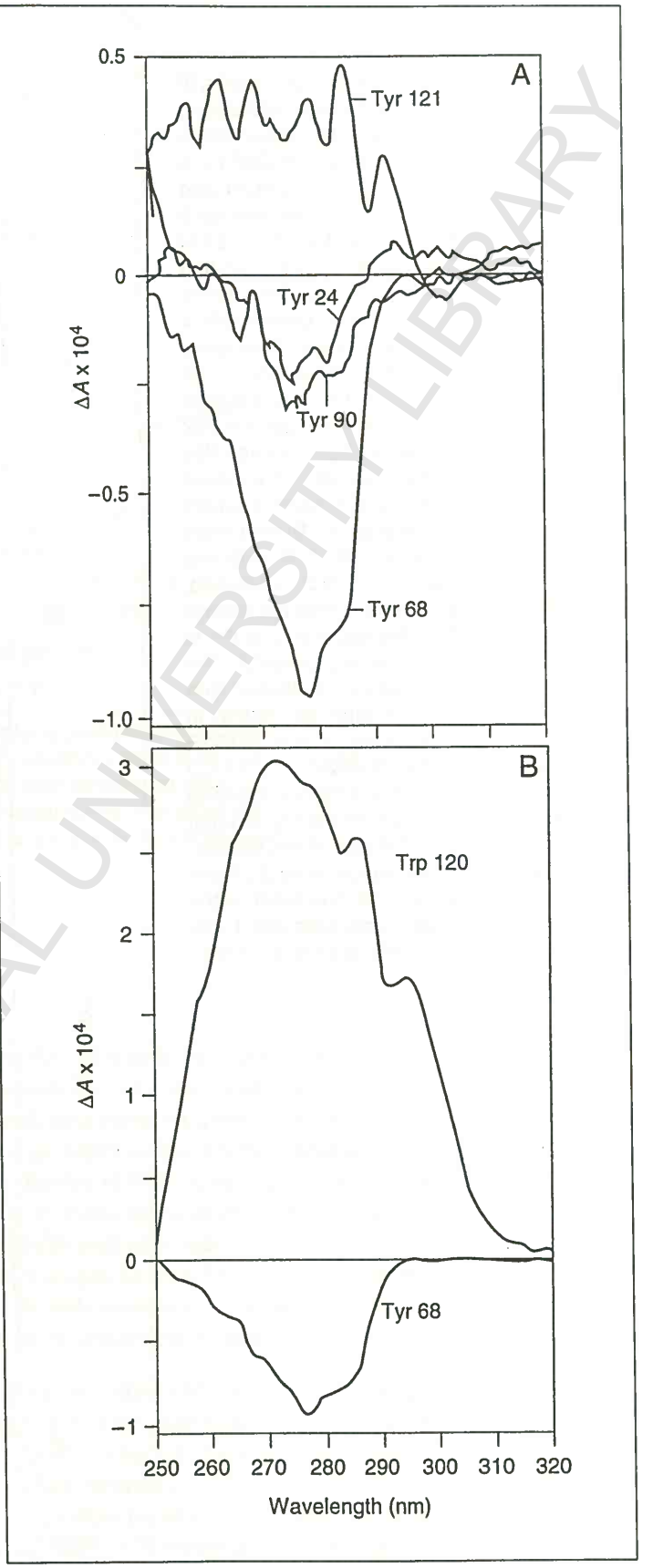


Figure 7.6.4 The contributions of different aromatic residues in a folded protein to circular dichroism. Interleukin 1β contains one tryptophan and four tyrosine residues. Their individual contributions have been resolved using site-directed mutagenesis. (**A**) The near-UV CD spectra of the tyrosine residues show that, although Tyr 68 makes a significant contribution to the negative ellipticity, indicating its buried, asymmetric environment, the contributions resulting from residues 24 and 90 are little different from that of free tyrosine. The derived spectrum of Tyr 121 is interesting in that it clearly contains a contribution arising from tryptophan (Fig. 7.6.2), indicating its close interaction with Trp 120. (**B**) The contribution from the tryptophan is seen to dominate, reflecting its nonpolar environment. The intensity of the protein spectrum is less intense than it might otherwise be, because of the opposing signs of the individual spectra. Spectra were measured using a 1-nm bandwidth (Craig et al., 1989).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

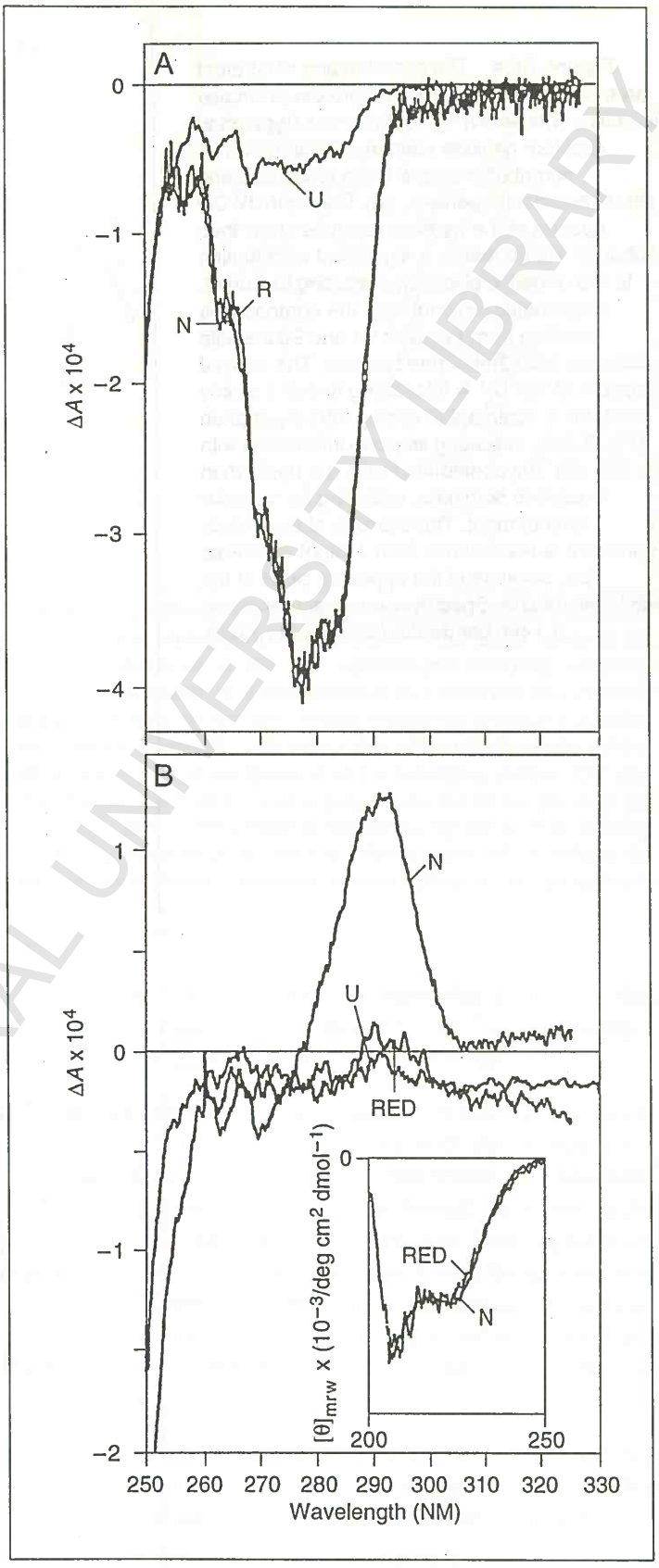


Figure 7.6.5 The near-UV CD spectra of granulocyte/macrophage-colony stimulating factor (GM-CSF). (**A**) Comparison of the spectrum of the native (N) murine GM-CSF (with one tryptophan residue, which is exposed, and four tyrosine residues, 2.5 of which are exposed) with those for the protein unfolded in 3.25 M guanidine·HCl (U) and refolded from this solution (R) by dialysis against 3 M urea. The intense negative spectrum of the folded protein, with a maximum at 278 nm, is characteristic of tyrosine, with contributions from phenylalanine. The intensity is probably a result of the interaction of Tyr 81 and Tyr 82. The absence of a band at 292 nm shows that the Trp 14 makes little contribution, as expected from its exposed location. Refolding restores the detailed tertiary structure of the native protein. (**B**) Reduction (RED) of native human GM-CSF (with two tryphtophan residues, one of which is exposed, and two tyrosine residues, which are buried) reduces the ellipticity of the native protein (N) to that of the fully unfolded protein (U). The secondary structure is however, unaffected (see inset for far-UV CD spectra). Notice that, in contrast to the murine protein, human GMCSF exhibits a less intense, but positive, near-UV CD band at 290 nm with little fine structure, implying a contribution from a ^1L_a transition of Trp 123 and indicating the buried nature of this residue. Spectra were measured using 0.01-mm path-length cells and 2-nm bandwidth in the far-UV, and 10-mm path-length cells with 1-mm bandwidth in the near-UV (Wingfield et al., 1988).

ACU IASI / CENTRAL UNIVERSITY LIBRARY

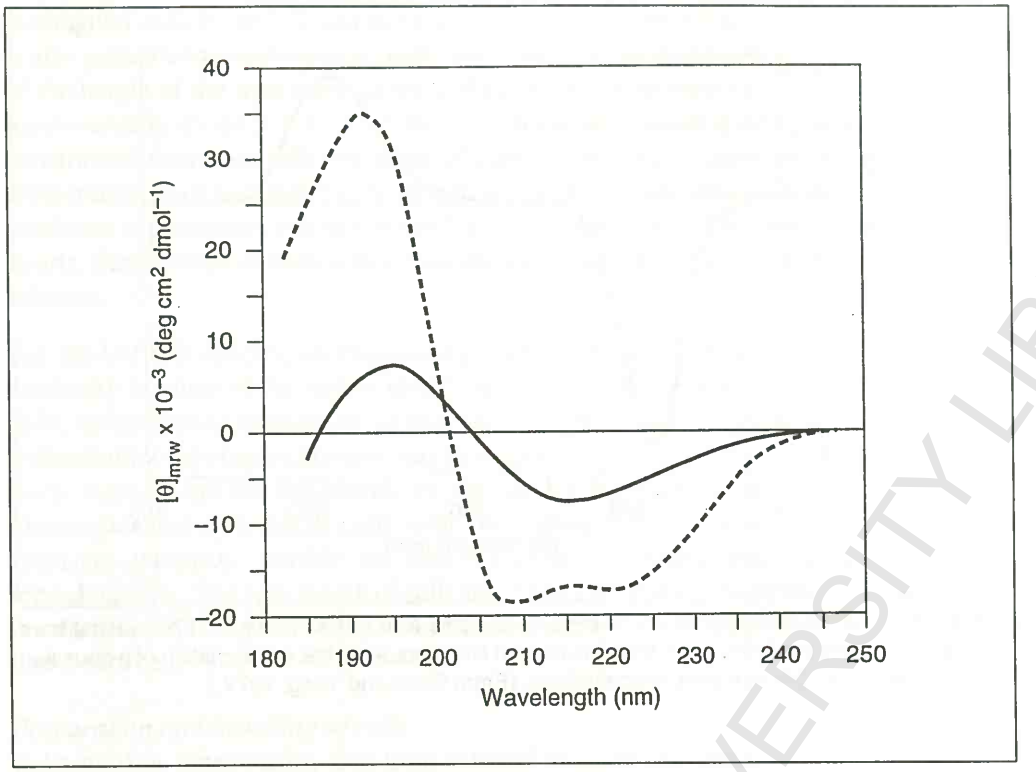


Figure 7.6.6 Characteristic far-UV CD spectra of proteins containing high contents of helix (dotted line) and β-sheet (solid line) structures. Protein concentrations were 0.36 mg/ml human serum albumin and 0.03 mg/ml immunoglobulin G. Spectra were determined using a 1-mm path-length cell. Each spectrum is the smoothed average of four repeat scans, measured with 1-nm bandwidth, reading ellipticity every 1 nm with an averaging time of 5 sec—i.e., at a scan speed of 12 nm/min.

STRATEGIC PLANNING

CD Spectrometer

CD spectra are measured using a CD spectrometer, of which there are three different makes on the market (AVIV Associates, Instruments SA, and Jasco; see *SUPPLIERS APPENDIX*). These instruments are fitted for control, data collection, and data handling by computer. The principles of how they function are generally outlined in the manufacturers' manuals and a good general account is given in Bayley (1980). A supply of high-purity nitrogen is essential to displace oxygen, in order to avoid degradation of the mirrors by ozone generated by the high-power xenon sources as well as to reduce absorbance from oxygen bands below 200 nm (see discussion of Nitrogen Supply, below). A satisfactory means of thermostatting cells is essential for work with proteins (see discussion of Cells, below), using a water-circulation system or the more convenient Peltier thermostat.

The CD spectrometer is usually required to work near the limits of sensitivity—e.g., reading ΔA values of $\leq 10^{-4}$ at a total absorbance of 1. Thus, particular care needs to be taken with cleanliness and orientation of cells and with settings of scan rate, time constant, and bandwidth. It is also important, especially when recording far-UV spectra, that the lamp is not old and that the mirrors are not clouded from radiation and traces of ozone. Because the spectrometer is a single-beam instrument, it is essential always to watch carefully for evidence of instrumental drift during measurements of sample and baseline.

BOSTON UNIVERSITY LIBRARY

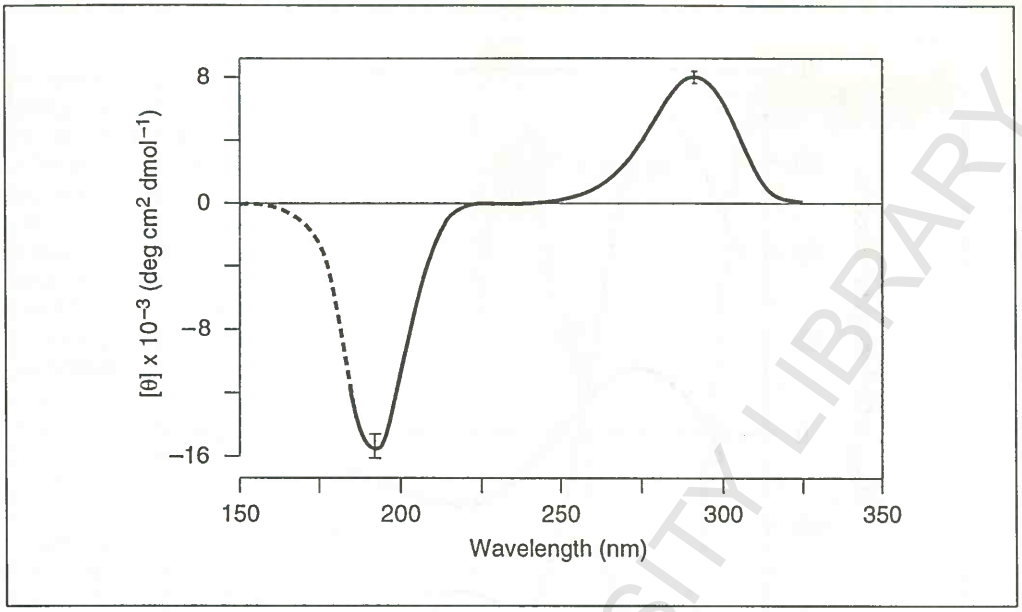


Figure 7.6.7 CD spectrum of D-(+)-10-camphorsulfonic acid (CSA) in water. The vertical bars represent variations of ±1.5% and ± 5%. The broken line represents the extrapolation of a gaussian band. Commercial CSA was twice recrystallized. (From Chen and Yang, 1977.)

Calibration of the Spectrometer

Depending on the make of the spectrometer, wavelength calibration may require the use of a mercury lamp and the services of an engineer, or the use of a holmium oxide solution with a simple computer-operated adjustment. It is unlikely to vary significantly during the lifetime of a xenon lamp. Calibration of ellipticity, on the other hand, must be carried out regularly—at least monthly, if the machine is in regular use—using a purified compound of known absorbance and ellipticity. D-(+)-10-camphorsulfonic acid (CSA) is the substance of choice (Fig. 7.6.7), but care must be taken to purify it from optically active impurities. Crystallization from ethyl acetate is recommended. CSA is hygroscopic; hence, although a stable monohydrate can be prepared (Yang et al., 1986), the concentration of the aqueous solution (~1 mg/ml) should be determined from $A_{285}^{1\ mg/ml} = 0.149$ using a high-quality spectrophotometer. Calibration of the spectrometer is carried out according to the maker's instructions, using $\theta = 33.5$ mdeg at 290.5 nm for a solution of 1.00 mg/ml anhydrous CSA in a 1-mm path-length cell. CSA exhibits a second peak at 192.5 nm where $\theta \approx -67.0$ mdeg for the same solution. If the value of the ratio of ellipticity at the two wavelengths falls below 2.0, this indicates a loss of performance of the spectrometer. Epiandrosterone, at 304 nm, and D-(−)-pantoyllactone, at 219 nm, have also been used for calibration (Schmid, 1989). The peaks of the spectra scanned for calibration are sufficiently sharp to provide a check on the wavelength calibration.

Cells

Choice of cells

Quartz windows can exhibit dichroism, which is due to strain either remaining after the annealing process during manufacture or induced by distortion caused by heating or pressure. It is important for measurement of CD that the optical faces of the quartz QS cuvettes have low strain and exhibit low dichroism. In the near-UV region—where path lengths ≥5 mm are required—rectangular cells, in their standard, semimicro, or micro forms, are normally used. These are not only more economical with regard to protein than the cylindrical type, but allow mixing to be carried out in the cell. Testing a number of

ACU IASI PERPUSTAKAAN UNIVERSITY LIBRARY

rectangular cells in the CD spectrometer usually yields one or two with low dichroism; it is also possible to purchase cells with certificates of low dichroism. Careful observation of the height of the light beam at the cell allows the minimum volume of solution to be used—usually <1 ml in a 10 × 10–mm cell. Because standard quartz (Suprasil) cells are constructed from two different types of glass, differential expansion of the faces with temperature leads to a significant change in cell dichroism. At least for temperature-dependence experiments, it is recommended to use cells with all four faces made from fused quartz. Semimicro cuvettes with 1-cm path length are available; these require ~300 µl of solution.

For far-UV CD spectra, short-path-length cells—from 1 mm down to 0.01 mm—are necessary because of the higher absorbance and ellipticity in this region. Demountable cells, cylindrical or rectangular, of this size are easily cleaned, but require practice to fill without allowing evaporation and may be subject to solvent loss over long periods of time. Such leakage can be minimized by the use of Parafilm on the edges of the cells. Quartz-jacketed cylindrical cells with two filling holes, available from Hellma (see *SUPPLIERS APPENDIX*), provide the best economy of protein, temperature control, and reproducibility. The path length of cells should be checked by absorbance measurements. A solution of 300 mg/liter of potassium dichromate in 0.05 M KOH gives an A_{280} of 0.705 in a 1-mm cell.

B. C. U. „M. EMINESCU" IAŞI
C H I M I E

Preparation and handling of cells

Cells must be reproducibly free from external and internal contamination. Not only will contamination make a contribution to the measured ellipticity, but dirty surfaces cause bubbles to be trapped in short-path-length cells, leading to additional artifacts. Cleaning should always be repeated if there are any signs of "tailing." The short-path-length cells, being inaccessible to mechanical cleaning, have to rely on effective soaking and rinsing. Cuvettes must be thoroughly and reproducibly cleaned by soaking in a proprietary cuvette cleaning fluid such as Hellmanex II (Hellma), an alkaline liquid concentrate, or RBS-35 Detergent Concentrate (Pierce), following the makers' instructions. The makers of Hellma cuvettes caution against the use of ultrasonics as a means of boosting cleaning efficiency. 50% nitric acid is also effective, but requires care and a fume hood. Detergents frequently contain fluorescent material and, if used, should be removed from cuvettes with acid (2M HCl) followed by water. After the final rinse with copious amounts of high-quality distilled water, the cells are best dried using suction through a Pasteur pipet protected with a small length of fine plastic tubing. This is safer than using acetone or alcohol, which can leach out fats from fingers and deposit them on the cuvette. The faces of cuvettes must not be touched after cleaning. Filling, emptying, and rinsing between measurements of spectra must be done using pipets, to avoid spillage onto the outer faces. A Pasteur pipet protected with fine plastic tubing is satisfactory. The baseline scans, with buffer in the cell, must be examined as a constant check on the reproducibility of the state of the cell.

Buffer Solutions

Buffers must be optically inactive and have the lowest possible absorbance in the desired spectral region. Most buffers in common use are essentially transparent above 230 nm, which is fine for near-UV CD. For far-UV CD using a 1-mm pathlength cell, 0.01 M sodium phosphate buffer (*APPENDIX 2E*), is generally acceptable as a solvent for most purposes, and will still have a reasonably low absorbance between 190 and 185 nm—the lower the pH, the more transparent the buffer. NaCl should be avoided in the far-UV range. Solutions of 0.01 M $NaClO_4$, NaF, KF, or boric acid are all transparent down to the present range of commercial instruments, but are not frequently used except for vacuum UV measurements. The absorbance of 20 mM Tris·Cl is acceptable down to 190 nm. Buffers

in common use, together with their absorbances, are listed by Schmid (1989). Urea or guanidine·HCl solutions (high-purity spectroscopic grades must be used) absorb very strongly below 220 nm; even in the shortest-path-length cells, they are opaque below 205 nm.

Bubbles form readily in the cell and can distort the CD. It is important therefore to degas the buffer before filling the cell by applying vacuum in a suitable container, such as a Buchner funnel, and swirling.

Clarification of Solutions

Protein solutions must be freed from the turbidity that results from dust and aggregated protein. Turbidity will reduce the sensitivity of the measurements and can also cause artifacts in both the intensity and the shape of the CD spectrum, particularly in the far-UV (Swords and Wallace, 1993). Many proteins can be filtered, usually as a stock solution, using a 0.22-μm filter (ideally of the grade special for proteins) fitted to a syringe. If the protein is adsorbed by filters, the solution may be clarified by centrifugation; 10 min in a refrigerated microcentrifuge is convenient and usually effective. Note that the protein concentration must be measured *after* clarification. Buffer solutions are filtered on a larger scale, using a 0.22-μm filter.

It is essential to monitor clarification if artifacts are to be avoided. Indicators of adequate clarification that should be adopted for spectroscopic characterization are available from the procedures basic to the measurement of CD. First, determination of protein concentration requires an absorbance spectrum to be recorded on a good quality spectrophotometer from 240 to 350 nm. Aromatic amino acid residues do not absorb above 320 nm, so the spectrum between 320 and 350 nm should be only marginally above baseline. The presence of turbidity will result in a sloping spectrum in this region. From the absorbance spectrum obtained in this manner, the ratio A_{max}/A_{min} is calculated, where A_{min} is the absorbance in the trough in the region of 240 to 250 nm. This ratio will be specific for a given protein and should be measured for a sample that has been carefully and reproducibly clarified. Because of the strong wavelength dependence of scattering, the ratio is a very sensitive indicator, decreasing sharply with turbidity, and can be used as a semi-quantitative check on the clarity of subsequent solutions of the protein.

For membrane proteins in membrane fragments, the intrinsic turbidity offers special problems, which should be addressed by instrumental modifications. A useful discussion of the problems involved can be found in Swords and Wallace (1993).

Nitrogen Supply

Oxygen-free nitrogen can be obtained from cylinders, but must be filtered to remove oil contaminants. Liquid nitrogen provides a better supply of gas, but impurities accumulate at the bottom of the Dewar flask, so that the bottom 10% should not be used (Johnson, 1990). Accessories to monitor and purify nitrogen gas are available commercially.

BASIC PROTOCOL

RECORDING A CD SPECTRUM

Simple comparison of the CD spectra of a protein in buffer and in 6 M guanidine·HCl will immediately tell whether the protein has achieved some degree of folding. However, for the purpose of determining whether a sample of recombinant protein is correctly folded, the CD spectrum can provide a complex fingerprint that can be compared with that of a specimen that has been well authenticated as being native and functional. This constitutes an empirical, but powerful means of characterization, with the near-UV spectrum usually being more sensitive than its far-UV counterpart. There are, for example,

PCU IASI - CENTRAL UNIVERSITY LIBRARY

many well-documented cases where partially unfolded proteins exhibit a major change in the aromatic region accompanied by relatively minor changes in the far-UV CD (e.g., see Fig. 7.6.5)—the classical molten globule being an important example (Ptitsyn, 1992). For these reasons, it is important to obtain good quality near-UV spectra, paying particular attention to details of any fine structure that may exist.

Materials

CD spectrometer (calibrated), nitrogen supply, cells, buffer solution, and clarified protein solution for analysis (see Strategic Planning)

1. Set up the CD spectrometer by purging the optics with nitrogen in the manner designated in the manual, turning on the cooling water, and finally switching on the lamp. Allow to warm up for the recommended period of time, usually 30 min. Regulate the thermostat system to the desired temperature.

   Particular attention must be paid to the procedure in the manual for switching on and warming up the spectrometer. Failure to purge with nitrogen can lead to rapid and irreversible degradation of the optics.

   CD spectra, particularly in the near-UV (see Support Protocol 1) reflect the dynamics of the chromophore and may therefore show dependence on temperature. It is important to stabilize the temperature reproducibly. Accurate temperature control is particularly important in denaturation experiments to determine protein stability.

2. Enter the settings required for the scan.

   Careful optimization of the settings is necessary to achieve the lowest-noise spectra consistent with total time spent in data collection. To check that optimal settings have been achieved, a small sample region of the spectrum can be scanned.

   Table 7.6.1 is a rough guide for scanning a protein solution with A_{280} = ~1. Settings will differ according to the sample and the spectrometer.

   a. Set wavelength.

   For near-UV spectra, the scan should be from 240 to 340 nm, in order to cover contributions from tryptophan, tyrosine and phenylalanine. Although little or no dichroism will be found above 320 nm, it is advisable to extend the scan to 340 nm to ensure that both sample and baseline scans coincide at zero ellipticity in the region where there is no dichroism. This provides assurance that the cell and buffer are optically identical for the two scans (see Fig. 7.6.8).

   For far-UV spectra, it is usual to scan between 250 nm—where the ellipticities of the baseline and sample effectively coincide, given the scale of ellipticity being used—and

Table 7.6.1 Rough Guide for Scanning a Protein Solution with A_{280} = ~1

	Far-UV	Near-UV
Cell path length (cm)	0.01–0.05	1
Wavelength (nm)	250–180	340–250
Bandwidth (nm)	0.5–1	0.2–0.5
Averaging time (sec)	1	1–5
or		
Time constant (sec)	2–8	8
Repeat scans	1–2	2–4
Step size (nm)	0.5–1	0.2–0.5
Resulting time per scan (min)	70–140	3–37

BCU LRC CENTRAL UNIVERSITY LIBRARY

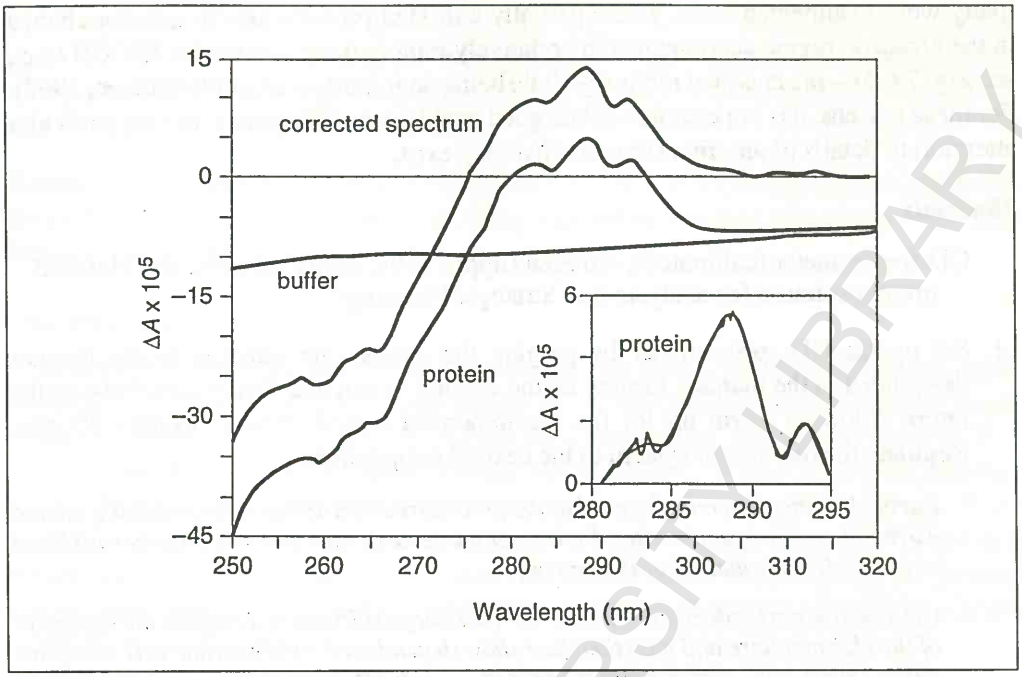


Figure 7.6.8 Obtaining the corrected near-UV CD spectrum for hen egg white lysozyme. The protein and baseline spectra were collected using a 10-mm cylindrical cell and 0.5 mg/ml protein in 0.067 M phosphate buffer, pH 6.0. Instrument settings were 1-nm bandwidth, 0.2-nm step size, scan speed 2 nm/min, time constant 8 sec (scan speed × time constant = 0.27 nm). Protein solution and buffer were scanned once each. The spectra were smoothed, a sample of the fit being shown in the inset. Reproducibility of the instrument and of the state of the cell are demonstrated by the coincidence of the ellipticity above 300 nm. The corrected spectrum was obtained by subtraction, using the instrument software.

as near to 180 nm as the absorbance of the solution will permit. The lower limit may be dictated by the absorbance of the sample or buffer, as indicated by the value of the dynode voltage (see step 2e).

b. Set bandwidth.

The radiation passing through the sample comprises a spread of intensities distributed around the selected wavelength (Fig. 7.6.9). The bandwidth (in nm) of this spread is determined by the spectrometer slit width (in mm) and is the width of the distribution at half the maximum intensity. For proteins exhibiting broad CD bands, including most far-UV and many near-UV spectra, a wide setting (2 nm) will have little effect on the shape or intensity of the spectrum. For bands with fine structure, including some near-UV spectra, a more narrow bandwidth is necessary to avoid smoothing of narrow peaks and reduction in peak intensity. The largest bandwidth that causes no distortion of the spectrum should be used, in order to maximize the signal-to-noise ratio. To optimize the slit-width setting, preliminary scans should be made at a series of slit widths over a critical region of the spectrum.

c. Set averaging time or time constant, scan speed, and number of accumulations.

The time constant determines the time that the spectrometer takes to respond to changes in signal. The reason for setting it is to obtain the best signal-to-noise ratio consistent with the time spent scanning. When the signal is strong, the time constant need only be small. The appropriate time constant for a signal of 500 mdeg is 0.25, that for a signal of 100 mdeg is 0.25 to 0.5, that for a signal of 10 mdeg is 0.5 to 0.8, and that for a signal of 2 mdeg is 2 to 30.

The faster the scan speed (with economies in operator time), the smaller the time constant must be in order to allow the instrument to respond fully to changes in ellipticity. However, less time is then available for collecting data at a given wave-

PCU IASLIC LIBRARY UNIVERSITY LIBRARY

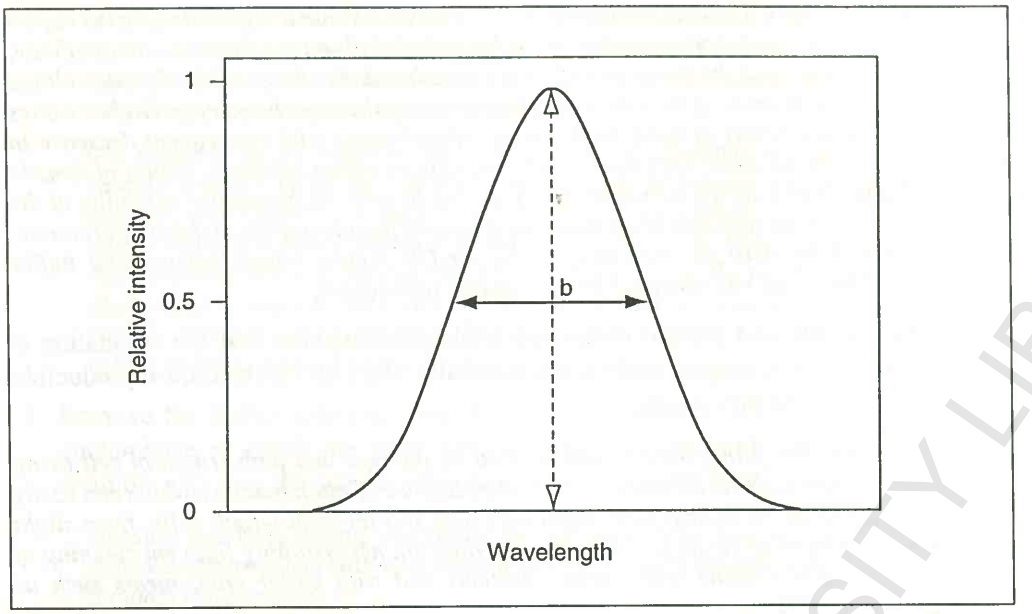


Figure 7.6.9 The distribution of intensity of radiation passing through the sample cell. Radiation emerging from a monochromator set at a wavelength corresponding to the peak maximum will contain components at other wavelengths, with intensities described by the curve. The bandwidth (b) is the width of the distribution curve at half the peak height. Its magnitude will depend on the exit slit width. The setting of the slit width is usually programmed to change during scanning, so as to give a constant bandwidth and hence constant spectral resolution.

length, so the noise is greater. Scan speed and time constant are thus interdependent, and their product should not exceed 0.33 nm. For example, a time constant of 2 sec is recommended for a scan speed of 10 nm/min. Some instruments, in the wavelength scan mode, specify an averaging time, which takes the place of a time constant setting; this is simply the time spent collecting data at each wavelength.

The signal-to-noise ratio can be enhanced either by increasing the time constant or averaging time or by accumulation of a number of scans. The latter technique is preferable to using long scan times because the effect of instrument drift will be to distort the spectrum from a single long scan, whereas the average of quicker, repeat scans will lead to a more uniformly shifted spectrum. The latter can then be corrected using the average of buffer baselines measured before and after the sample scans.

The problem of optimizing signal-to-noise ratio is greater in the case of samples possessing a low ellipticity. An all-β protein, for example, will demand more repeat scans than an all-helix protein at the same concentration.

There are realistic time limits to the degree of enhancement of signal-to-noise ratio, as the latter is proportional to the square root of both the time constant and the number of scans (see Yang et al., 1986).

d. Set step size.

The step size (in nm) between successive data-collection wavelengths, together with the scan speed, defines the number of data points collected per nm. Spectra with more fine structure require higher resolution. For far-UV spectra, which are smooth and relatively featureless, a step size of 0.5 nm is adequate. For some, near-UV spectra, but not all, a resolution of 0.2 or 0.1 nm is advisable. For a given scan speed, the greater resolution gained by using a small step size will result in more noise, so it is more economical of time to use as large a step size as possible.

e. Note dynode voltage.

Dynode voltage cannot be preset, but its value can be controlled within limits by adjusting the protein concentration and/or cell pathlength, so it is appropriate to

BOSTON COLLEGE UNIVERSITY LIBRARY

consider it here. The photomultiplier detector operates with a high voltage in the region of 1000 V across the anode and cathode. When light falls on the photosensitive cathode, current flows and the potential difference is reduced. Readings of the dynode voltage thus give a measure of the intensity of light passing through the sample. Higher values indicate low levels of light (high sample absorbance) with consequent decrease in signal-to noise-ratio; very high values indicate incorrect readings. Values of dynode voltage should always be noted during a scan in order to assess the reliability of the ellipticity measurements. Most instruments record this parameter as a matter of course. Special care must be exercised in the far-UV region, where sample and buffer absorbance can rise sharply with decreasing wavelength.

3. Fill the cuvette and place it in the cell holder, making sure that the orientation is reproducible both longitudinally and rotationally. Mark the cell to allow reproducible orientation in the light beam.

   Techniques for filling the cell will depend on the type and path length of cell being used. It is essential in all cases to avoid the presence of small bubbles, which can easily be trapped at the optical face, especially with shorter-path-length cells. Even slight contamination of the faces will make this difficult, necessitating frequent cleaning of the cell after contact with certain proteins and with buffer components such as guanidine·HCl.

   For rectangular-type cuvettes, the volumes of solution required are approximately 2.5 ml (for 1-cm standard), 1 ml (for 1-cm semimicro), 300 µl (for 0.1-cm standard) and 25 µl (for 0.01-cm demountable). Volumes for cylindrical-type cells are somewhat larger, depending on the type.

   Choice of path length will depend on the available concentration of the protein solution. If the absorbance of the buffer is unavoidably high, it will be advantageous to use a shorter path length and a higher protein concentration. As a general guide, for a protein solution with $A_{280} = 1$, path lengths of 0.01 to 0.05 cm for the far-UV and 1 cm for the near-UV make suitable starting points.

   Rectangular cells have the advantages over cylindrical cells of having smaller sample volume and better temperature control with available cell holders. Good temperature control with cylindrical cells requires the use of jacketed cells. For shorter (<0.1-cm) path lengths, where rectangular cells are currently demountable, the rigid cylindrical cells have the advantage of avoiding leakage over prolonged periods of time.

4. Scan the baseline using the same cell and buffer as required for the sample, as well as the same instrument settings.

   Preliminary scans of the samples should be carried out to assess whether valid measurements can be made in the required wavelength region or whether it is necessary to adjust the sample concentration or the cell path length. The wavelength is lowered, noting the point at which the photomultiplier voltage exceeds the value corresponding to low light intensity—i.e., to an absorbance of 1 to 1.5. Experience will indicate the value above which noise becomes excessive. It is helpful, however, to get a more objective assessment by noting the voltage at a series of low wavelengths and then measuring the actual absorbance of the same cell using a quality spectrophotometer.

   When recording spectra for a series of samples, buffer scans should be recorded at least twice (at the beginning and end of the series) in order to check for instrument drift, especially when the lamp is beginning to age.

   A convenient procedure when scanning a number of samples is to scan air blanks—i.e., with no cell in the light path—at appropriate intervals between samples. This saves additional cell filling, washing, and drying operations. A single scan (S) of the cell containing buffer accompanied by a scan of the air blank then allows correction of all the protein spectra for solvent baseline, according to the expression:

$$\{S_{protein} - S_{air}\} - \{S_{solvent} - S_{air}\} = \{S_{protein} - S_{solvent}\}.$$

ACU IASI CENTRAL UNIVERSITY LIBRARY

Ideally, the buffer scan should be run before the protein sample, as this requires less rigorous washing of the cell between the two runs and hence better assurance that the optical surfaces will be identical for both. This however requires prior knowledge of the settings appropriate to the particular sample.

It is tempting to use water in place of buffer on the grounds that both are optically inactive and, for water, the cell needs only to be dried between scans. This is unwise, however, particularly for buffers with higher concentrations of solutes and hence higher refractive index. Because in most instruments the light beam is not parallel when passing through the cell, the area and position of the photomultiplier surface illuminated is likely to vary with the refractive index of the solution, which may give slightly different voltage readings. This may result in significant errors when the signal is low, as in the case of near-UV CD.

5. Remove the buffer solution, rinse the cell with water and dry, taking care not to contaminate or touch the outer surfaces, refill the cell with the clarified sample solution, and scan the sample using the same instrument settings as for the buffer.

Because washing, drying, and filling the cell takes time, it is advisable to work with two cells alternately (both with solvent blanks properly determined) when processing a number of samples.

The protein concentration should be as high as possible, consistent with the absorbance in the cell not exceeding 1.0 to 1.5, and must be accurately determined by UV absorbance. Values of mean residue ellipticity and secondary structure content results are both affected by uncertainties in protein concentration.

In spectral regions such as the far-UV, where absorbance varies markedly with wavelength, the spectrum may have to be recorded in overlapping regions, at appropriate concentrations or cell path lengths.

A good test for the absence of artifacts that can arise at higher absorbances is to run a blank with a non–optically active molecule, such as 3-methylindole or a DL–amino acid, at the same absorbance as the test samples. The scan should be indistinguishable from the solvent blank.

6. Subtract the baseline from the sample scan.

The ellipticity of the spectrum corrected for baseline should be close to zero above 320 nm for near-UV and above 250 nm for far-UV spectra (see Fig. 7.6.6 and Fig. 7.6.8). Significant differences usually indicate problems with reproducibility of the buffer or cell and must be remedied for quantitative characterization of samples.

Depending on the level of noise, spectra and baselines can be smoothed using the instrument software. Smoothed and raw spectra should be compared directly by overlaying on the computer screen in order to avoid the distortion that can arise from overenthusiastic smoothing.

INTERPRETATION OF NEAR-UV CD SPECTRA

SUPPORT PROTOCOL 1

CD spectra reflect, in different ways, the detailed conformational environment of individual chromophores in a protein. To the extent that the transitions for the different types of chromophore differ in wavelength, the spectra can be assigned to particular types of residue and, in the case of the peptide bond, to different conformations or types of secondary structure. A brief account of the interpretation of near- and far-UV spectra is therefore pertinent.

Near-UV CD spectra provide evidence of tertiary structure and dynamics. The CD of chromophores absorbing in the near-UV region is associated with the π–π^* electronic transitions. Relevant transitions for the aromatic amino acids are listed in Table 7.6.2. Tryptophan, tyrosine and phenylalanine possess partial or complete planes of symmetry, so that their dichroism arises almost entirely as a result of interaction of the transitions

BOURNEMOUTH UNIVERSITY LIBRARY

Table 7.6.2 Transitions of Aromatic Amino Acid Side Chains[a]

Amino acid	Electronic transitions	Peak wavelengths (nm)	$\Delta\varepsilon_{max}$ (M^{-1} cm^{-1})[b]
Phenylalanine	^1L_b	262, 268	±0.3
Tyrosine	^1L_a	far UV	
	^1L_b	277, 283	±2.0
		268, 289 (shoulders)	
Tryptophan	^1L_a	266	±3
		~295-303[c]	
	^1L_b	285, 292, 305[d]	±2.5

[a]Data from Strickland (1974).

[b]Estimated values for an isolated side chain.

[c]May be seen as a low intensity shoulder or peak from a Trp in a nonpolar environment.

[d]Unusual; observed with, e.g., lysozyme.

with neighboring groups. Dipole-dipole interactions, especially with other aromatic rings and with the π–π^* transition of peptide bonds, make a major contribution to CD and are relatively long-range, being effective up to ~1 nm. Electric-magnetic coupling also has an influence on the CD, but is very short-range and makes a relatively minor contribution. In contrast to the aromatic residues, the disulfide bond has no plane of symmetry and is, as a result, intrinsically optically active, particularly around a dihedral angle of 90°.

These interactions of the environment with transitions of aromatic residues result in dichroic bands that are either positive or negative so that statistically the more such residues there are, the less intense the resultant ellipticity is likely to be. An intense signal frequently indicates the presence of one particularly strong interaction that overrides other minor contributions—such as the strong band at 292 nm seen in aspartic proteinases (Fig. 7.6.10), and the contributions of the one tryptophan and four tyrosine residues to the CD spectrum of interleukin 1β (Fig. 7.6.4).

The ^1L_b transition for **phenylalanine** resembles that for benzene (Strickland, 1974), but the CD signal is of relatively low intensity, so that it is seen only as shoulders or minor peaks to the low-wavelength side of tyrosine and tryptophan bands (Fig. 7.6.10). As a result of the high degree of symmetry in the side chain, the presence of these peaks is an indication of a highly specific and well-packed environment and can act as a marker for native folding.

The ^1L_b transition for **tyrosine** (the ^1L_a transition lies in the far-UV) is more intense than that for phenylalanine and overlaps the two transitions for tryptophan (Strickland, 1974). It can result in a significant contribution to the near-UV CD (e.g., see Fig. 7.6.4 and Fig. 7.6.5) but is not always easy to assign in the presence of tryptophan. Hydrogen bonds to peptide carbonyl, carboxylate (C=O), and imidazole (–N=) groups contribute significantly to its CD, as do the presence of neighboring polar groups. The direction of the optical transition is in the plane of the ring and perpendicular to the axis of symmetry (Fig. 7.6.11) so that interaction with neighboring groups makes the CD very sensitive to rotation of the residue and to the dynamics of the protein in general.

The two transitions for **tryptophan** exhibit distinct features, which lead to quite different CD bands. The ^1L_a transition is broad, relatively featureless, and intense (Fig. 7.6.2). The ^1L_b transition is weaker but exhibits fine structure (Fig. 7.6.2) similar to that of tyrosine. It can be masked by or superimposed on the ^1L_a transition. Interactions similar to those listed for tyrosine strongly affect the peak wavelengths and intensities. The presence of

PCU IASI/CENTRAL UNIVERSITY LIBRARY

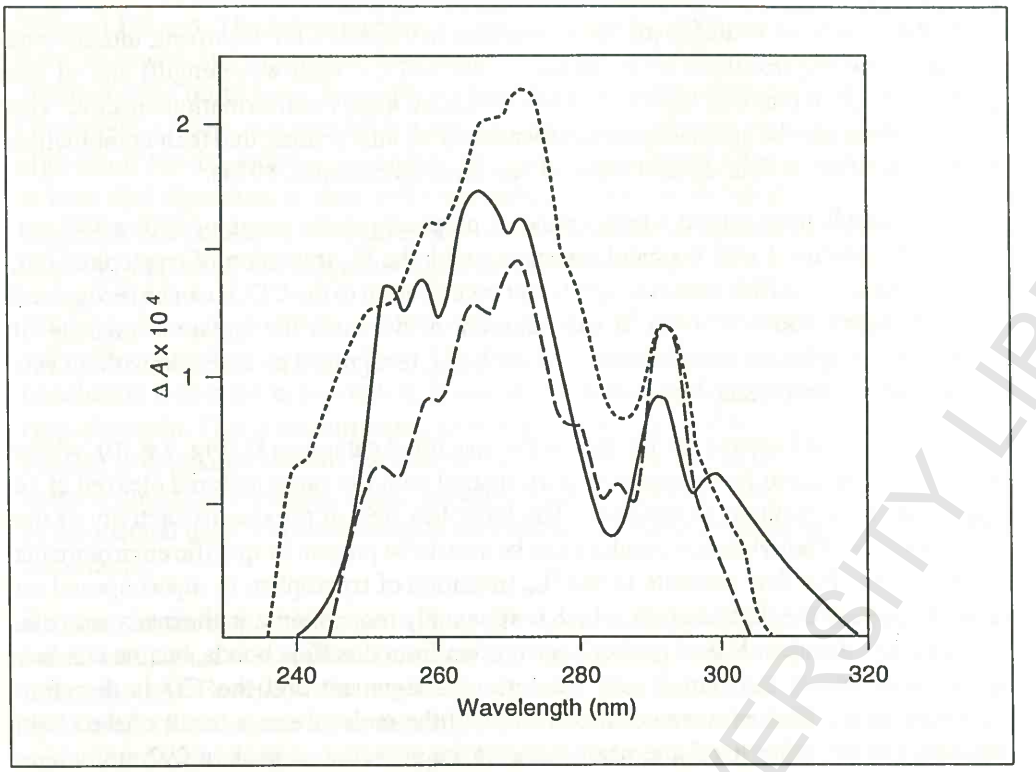


Figure 7.6.10 The near-UV CD spectra of cathepsin D. The solid line represents the porcine enzyme and the dashed line represents the bovine enzyme, each cleaved into two chains which remain associated; the dotted line represents the bovine intact enzyme. The spectra were recorded using 10-mm- and 20-mm-path-length cells, with enzyme concentrations of 0.8 to 0.4 g/liter. The far-UV CD and fluorescence spectra of the cleaved and intact enzymes are not significantly different (Pain et al., 1985).

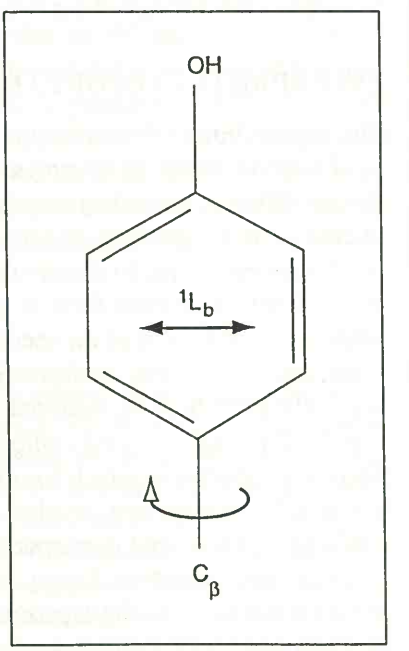


Figure 7.6.11 The near-UV absorbance transition dipole moment of the tyrosine side chain. The ^1L_b transition is seen to be perpendicular to the axis of rotation of the phenolic ring. Interactions of the transition with the environment, expressed as an enhanced CD band, will be sensitive to rotation of the side chain around this axis.

BCU IASI / GENERAL UNIVERSITY LIBRARY

bands above 285 nm is diagnostic for tryptophan in a specific environment; interactions that cause the ^1L_b transition to be shifted to the red (or high-wavelength) end of the spectrum result in bands as high as 310 nm which are highly conformation-specific. The two transitions may be affected quite independently by interactions, and their combination can thereby result in four distinct types of spectrum (Strickland, 1974).

Disulfide bonds may exhibit a broad band of ellipticity in the range of 240 to 350 nm. This can be confused with the band associated with the ^1L_a transition of tryptophan but, in cases where a disulfide makes a significant contribution to the CD, it can be recognized by its ellipticity above 320 nm. It can augment or diminish the apparent intensity of tryptophan and tyrosine contributions without being recognized as such and without fine structure necessarily being lost.

Some of the above features can be seen in the spectra of cathepsin D (Fig. 7.6.10), where the intact single-chain bovine enzyme is compared with the same material cleaved at an exposed loop but without dissociation. The latter has 50% of the specific activity of the intact molecule. Phenylalanine residues can be seen to be present in specific environments in both forms. The fine structure of the ^1L_b transition of tryptophan is superimposed on the broad peak of the ^1L_a transition, which is apparently more intense in the intact enzyme. Alternatively, there could be a greater contribution from disulfide bonds, but the absence of ellipticity above 320 nm favors the former assignment and the CD is therefore consistent with a limited increase in dynamics of the molecule as a result of the chain cleavage. On the other hand, the persistence of the pronounced peak at 292 nm, where the contribution of the ^1L_a transition is low, indicates that the interaction responsible for this band is unchanged by cleavage. Coupled with the fact that both far UV CD and fluorescence spectra are identical for the two forms, it is concluded that the overall conformations are structurally similar, differing mainly in their packing and dynamics. Comparison with the spectrum for porcine cathepsin D demonstrates the conformational similarity expected from sequence homology, as evidenced in particular by the persistence of the interaction responsible for the appearance of the ^1L_b band with its intense peak at 292 nm. Differences in the relative intensities of the minor bands provide a sensitive fingerprint for distinguishing two proteins with strong sequence homology.

SUPPORT PROTOCOL 2

INTERPRETATION OF FAR-UV CD SPECTRA

The peptide bond exhibits intense circular dichroism bands when it is located in regular, rigid conformations of the protein backbone chain. It is not surprising therefore to find that the different secondary structures—helix, pleated sheets, and β-turns—each give rise to characteristic spectra (see Fig. 7.6.6). In principle it should be possible, knowing the spectra for each type, to deconvolute the far-UV spectrum for a particular protein to give the proportions of each form of secondary structure in the folded conformation. Early work involved the use of the spectra of polyamino acids as models for α-helix, β-pleated sheet, and random coil conformations. These turned out to be inadequate models for the relatively short helices, distorted intramolecular β-sheets, and aperiodic—but far from statistically random—coil conformations found in the average globular protein. The alternative approach, which forms the basis of most estimations and avoids the need to define reference spectra, involves the use of a database of CD spectra (see Yang et al., 1986, for listings) and corresponding secondary structure contents derived from X ray structure determination. Linear combination of these spectra is carried out to give a resultant that best fits the experimental spectrum for the test protein. The two algorithms that have been most commonly used are those of Provencher and Glöckner (1981) and of Johnson and coworkers (Johnson, 1990). The former used a bank of 16 proteins with spectra measured between 240 and 190 nm. Using a different statistical approach, Johnson and coworkers provided an algorithm that uses 15 reference spectra measured between

PCU IASI CENTRAL UNIVERSITY LIBRARY

260 and 178 nm. The latter, together with a more recent development of the method, has been reviewed and compared with the Provencher and Glöckner algorithm (Johnson, 1990) and the underlying assumptions have been critically discussed (Manning, 1989). Greenfield (1996) has provided a useful and comprehensive review of all the main algorithms for determining secondary structure. Commercial CD spectrometers include at least one algorithm in their software, although it is worth installing the two original programs independently so that results from each can be compared. Copies are available from the authors of the programs and from Dr. Greenfield.

Attempts have been made to assess the success of different algorithms in deriving secondary structure content from CD spectra (see Yang et al., 1986). The general conclusion is that the helix content is usually estimated well, regardless of the structural type of protein. This is not surprising given the high intensity of the helix spectrum. β-sheet content is estimated well in the case of "all-β" proteins, but in mixed α and β proteins there is greater uncertainty. Some proteins, such as subtilisin, perform badly regardless of the method used. The estimation of the content of β-turns is similarly subject to greater uncertainty.

Reasons for these problems are various. First, estimations are dependent on the correct values of mean residue ellipticity, which depend critically on determination of the protein concentration. Second, it is reasonably claimed that the further down in wavelength spectra are recorded, the better the estimate of structure (Johnson, 1990). It is not possible with all proteins to obtain reliable spectra below 185 nm with commercial instruments—inability to use the appropriate buffer and intrinsic scattering from larger proteins are limiting factors, although loss of optimal instrument performance as a result of insufficient maintenance is frequently the main cause of the problem. Third, there is uncertainty about the degree of contribution of aromatic and cystine residues in the far-UV (see Chaffotte et al., 1992, for references and discussion). Aromatic residues are probably significant when the residues are located in clusters, and cystine residues when the dihedral angle is of the order of 90°. Fourth, criteria vary for defining the secondary structure content from the three-dimensional coordinates of structures.

In conclusion, the analysis of spectra properly recorded to 185 nm, or lower where possible, can give useful estimates of secondary structure content, but the content of turns and of β-structure should be interpreted with caution. In principle, Fourier transform infrared spectroscopy should provide better estimates of the latter. Secondly, when using the results of far-UV CD determination to characterize reproducibility of folding for different samples, it is important first to compare the spectra visually and to look for possible trends or factors that may explain small differences, rather than to rely solely on comparison of derived secondary structure contents.

COMMENTARY

Background Information

CD spectra in the near-UV are of much lower intensity than those in the far UV, in part as a result of the much lower concentration of the chromophore, but can be measured reproducibly with due care. They are very important in characterizing the degree and correctness of folding of recombinant proteins and their mutants.

Near-UV CD depends critically on the atomic environment and closeness of packing of the aromatic residues, including their accessibility to solvent. Although the structural information is limited strictly to the immediate neighborhood of the relatively few aromatic residues, the generally cooperative nature of protein conformation means that changes in aromatic CD frequently reflect more widespread changes in conformation and dynamics. A slight loosening of a protein structure may not result in any change in solvent accessibility to aromatic residues and therefore will not change the maximum fluorescence wavelength of tryptophan. Such a change may also be

RESTRICTED / CENTRAL UNIVERSITY LIBRARY

outside the resolution of X-ray crystallographic analysis. The resulting change in dynamics of the residue, and hence in the asymmetry and detailed interactions with the immediate protein environment, will however affect the near-UV CD spectrum, usually quite markedly (see Fig. 7.6.5 and Fig. 7.6.12).

Changes in the far-UV CD spectrum resulting from slight changes in conformation or partial unfolding will not in general be so marked as in the near-UV region, so that the value of the former for the detailed characterization of correct folding is not as great as that of the latter. It is, however, applicable to proteins with little or no aromatic content. The fact that the far-UV CD arises from the transitions of a single type of chromophore—i.e., the peptide bond—and that the backbone conforma-

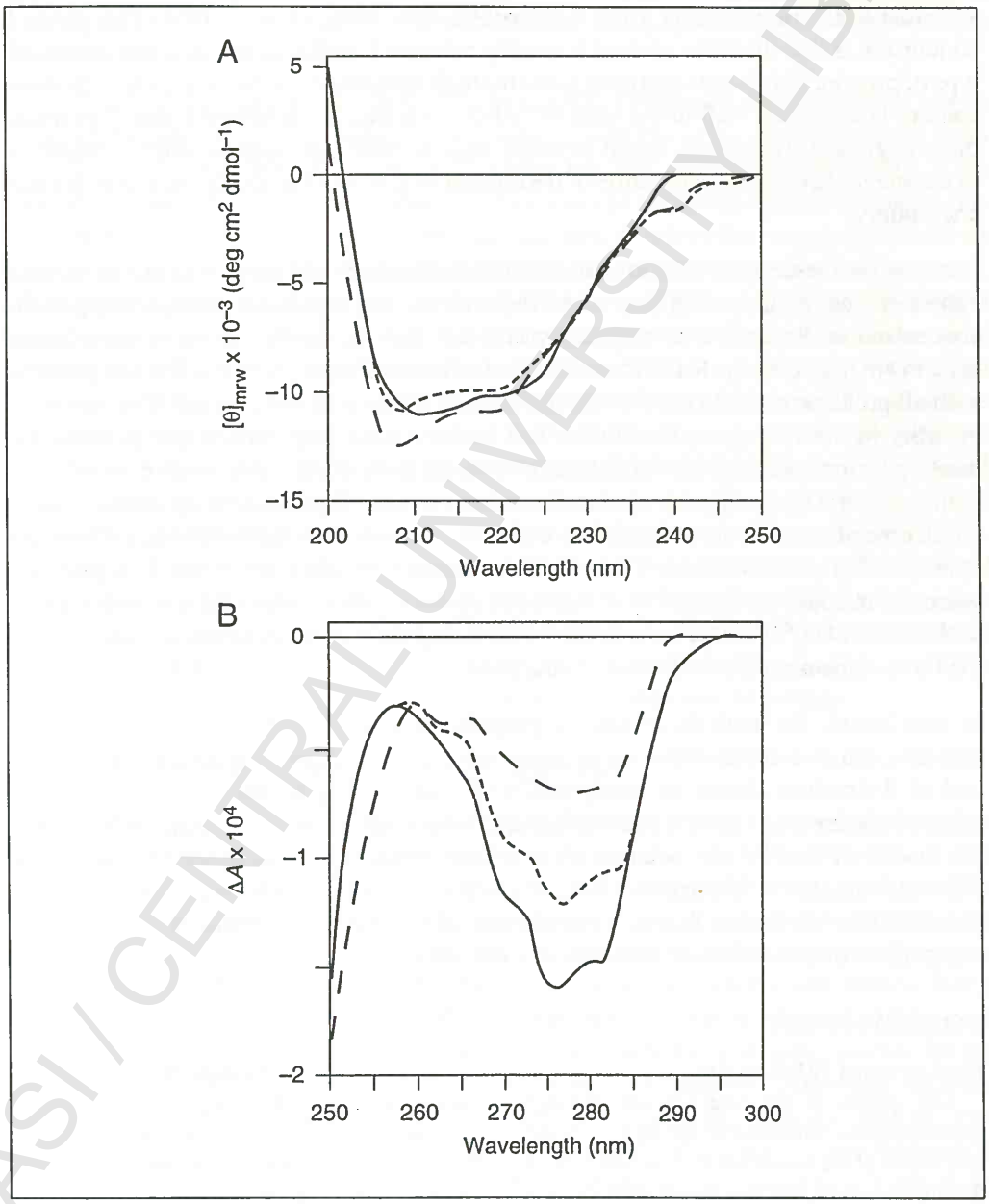


Figure 7.6.12 Effect of mutations detected by CD. The far-UV CD spectra (**A**) show that the secondary structure of β-lactamase PC1 (solid line) from *Staphylococcus aureus* is essentially unaffected by point mutations P2 (Thr 140→Ile; dashed line) and P54 (Asp 146→Asn; dotted line). The crystallographic structure of P54 (Herzberg et al., 1991) confirms that, apart from a loop region, the main body of the molecule that contains the thirteen tyrosine residues is very closely similar to that in the wild-type enzyme. The intensity of the tyrosine ellipticity (**B**) is, however, markedly decreased in each of the mutants, the lower thermodynamic stabilities of which support the interpretation of increased dynamics (Craig et al., 1985).

ACU IASI / CENTRAL UNIVERSITY LIBRARY

tion constitutes the main influence on its dichroism, means that this region is good for identifying family relationships between proteins.

CD in general is a useful tool for examining proteins made by site-directed mutagenesis for structure-function studies, where it is important to know to what extent the conformation has been affected by the mutation. The far-UV spectrum will tell whether the mutant has folded with a backbone conformation similar to that of the wild-type protein, whereas the near-UV spectrum will detect many minor changes in conformation or dynamics (Fig. 7.6.12).

CD complements fluorescence as a technique for characterizing the folded conformation in solution and for providing, in a single spectrum, a highly specific fingerprint for the native state. It requires only modest amounts of protein and measurements are relatively quick to make. Like fluorescence, its use as a comparative method requires a control spectrum from a carefully characterized sample of the natural or recombinant protein, recorded under defined and optimal conditions, using a spectrometer in good condition.

Critical Parameters

Many of the critical experimental parameters for CD spectrometry are the same as those for measuring fluorescence. Spectra will only be as reliable as the protein solution, the machine, and the operator. Cells must be scrupulously clean, solutions must be well-clarified, and buffers must have a low absorbance. In addition, not only must the protein concentration be known precisely but, together with the path length, must be chosen to optimize the absorbance in the region of interest. This will sometimes require recording a spectrum in overlapping sections. Preliminary experiments to optimize the instrumental parameters are important. Lastly, it is essential that the spectrometer be well-maintained and up to specification. The useful life for a lamp is on the order of 800 to 1000 hr, but performance in the far-UV may start to fall off before this.

Troubleshooting

The problem most frequently met (though a casual scan of the literature shows that it is not so frequently recognized) is the loss of light intensity at low wavelengths. Under such conditions, many spectrometers will produce traces that are smooth, giving the appearance of being valid spectra, but which are in fact artifacts. The example of such an artifact shown in Figure 7.6.13 provides a cautionary tale. In this case, the glutamine absorbs light in the far-UV on account of its side-chain amide group. The increasing concentration of glutamine results in a progressively longer wavelength at which light is cut off, to which this particular spectrometer responded by returning to zero dichroism, giving the impression of a peak with a maximum at increasing wavelengths. To avoid such artifacts, it is essential to monitor the photomultiplier high voltage during a scan. Similar problems can arise from using too high a protein concentration or too long a light path, from too much absorbance by the buffer, from turbidity in the protein solution, from a dirty cell, or—below 200 nm—from absorption by oxygen that results from incomplete purging with nitrogen.

A further problem that is sometimes encountered is that the spectrum obtained after baseline subtraction does not return to zero above 320 nm (in near-UV CD) or 250 nm (in far-UV CD). If, above these wavelengths, the spectrum runs parallel to the zero base, a provisional spectrum may be derived by subtracting the difference in ellipticity, as a constant, from the whole spectrum. For quantitative assessment of a sample, however, the spectrum should be redetermined, paying attention to the correct buffer blank, cell reproducibility, and instrumental drift (see Strategic Planning).

Care must be taken to rectify all of the above mentioned problems if meaningful spectra are to be obtained for reliable characterization.

Anticipated Results

Several examples of near- and far-UV CD spectra are given in the figures included in this unit. With current instruments the final spectra exhibit low noise and—provided that the instrument and sample parameters have been optimized—should be true and reproducible within the limits of protein-in-buffer absorbance of <1. The most frequent source of nonreproducibility of intensity in spectra is the difficulty of determining the protein concentration, particularly of larger proteins that scatter more and of those with a greater tendency to aggregate.

What constitutes a significant difference between two spectra? When the differences are small, the answer depends on sample preparation and sample stability as well as accuracy of concentration determination, identification of and compensation for drift in the spectrometer, correct baseline correction, absence of bubbles

BOLAS/BETRAN UNIVERSITY LIBRARY

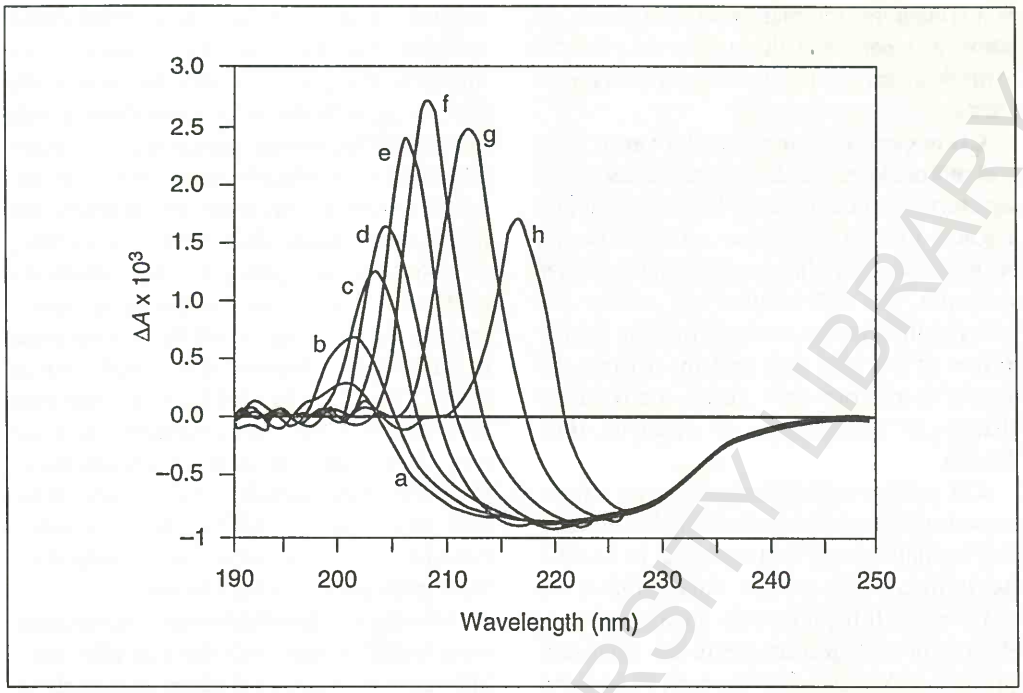


Figure 7.6.13 A convincing artifact. In an attempt to study the conformational consequences of adding an acceptor, D-glutamine, to a DD-peptidase, far-UV CD spectra were recorded for the enzyme in the presence of increasing concentrations of glutamine: a = 0 mM, b = 3 mM , c = 7 mM, d = 10 mM, e = 20 mM, f = 30 mM, g = 50 mM, and h = 95 mM. The enzyme concentration was 0.1 mg/ml, equivalent to ~10^{-3} M peptide bond, in 10 mM sodium phosphate pH 7.2. A 2-mm cell path length was used.

in the sample, reproducible cleanliness of the cuvette, and the level of general handling procedures. Ultimately, an assessment of significance depends on the experience, competence, and confidence of the operator.

As CD is very sensitive to small changes in conformation and dynamics, wild-type and mutant proteins will frequently give spectra that differ slightly in intensity rather than shape. If the dihedral angle for a disulfide bond is such as to yield an intense contribution, small conformational perturbations in a mutant protein may further affect the spectrum. Hence, there is the need to examine proteins by a number of different spectroscopic, hydrodynamic, and functional probes in order to assess correct folding of mutant proteins.

Time Considerations

A single spectrum, with associated baseline, will require ~1 hr, including the time needed to fill and clean the cuvette. Up to an additional hour may be needed for optimizing instrument parameters if the sample is unfamiliar. For a number of samples, economies of time can be effected (see Basic Protocol).

Literature Cited

Bayley, P. M. 1980. Circular dichroism and optical rotation. *In* An Introduction to Spectroscopy for Biochemists (S.B.Brown, ed.) pp. 148-235. Academic Press, London.

Chaffotte, A.F., Guillou, Y., and Goldberg, M.E. 1992. Kinetic resolution of peptide bond and side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry* 31:9694-9702.

Chen, G.C. and Yang, J.T. 1977. Two-point calibration of circular dichrometer with d-10-camphorsulfonic acid. *Anal. Lett.* 10:1195-1207.

Craig, S., Hollecker, M., Creighton, T.E., and Pain, R.H. 1985. Single amino acid mutations block a late step in the folding of β-lactamase from *Staphylococcus aureus*. *J. Mol. Biol.* 185:681-687.

Craig, S., Pain, R.H., Schmeissner, U., Virden, R., and Wingfield, P.T. 1989. Determination of the contributions of individual aromatic residues to the CD spectrum of IL-1β using site directed mutagenesis. *Int. J. Peptide Protein Res.* 33: 256-262.

Greenfield, N.J. 1996. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.* 235:1-10.

Herzberg, O., Kapadia, G., Blanco, B., Smith, T.S., and Coulson, A. 1991. Structural basis for the inactivation of the P54 mutant of β-lactamase

ACU IAS GEORGIA UNIVERSITY LIBRARY

from *Staphylococcus aureus* PC1. *Biochemistry* 30:9503-9509.

Holzwarth, G. and Doty, P. 1965. The ultraviolet circular dichroism of proteins. *J. Am. Chem. Soc.* 87:218-228.

Johnson, W.C. 1990. Protein secondary structure and circular dichroism: a practical guide. *Proteins Struct. Funct. and Genet.* 7:205-214.

Manning, M.C. 1989. Underlying assumptions in the estimation of secondary structure content in proteins by circular dichroism spectroscopy - a critical review. *J. Pharmacol. Biomed. Anal.* 7:1103-1119.

Pain, R.H., Lah, T., and Turk, V. 1985. Conformation and processing of cathepsin D. *Biosci. Rep.* 5:957-967.

Provencher, S. W. and Glöckner, J. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 20:33-37.

Ptitsyn, O.B. 1992. The molten globule state. *In* Protein Folding (T.E. Creighton, ed.) pp. 243-300. W.H. Freeman, New York.

Schmid, F.X. 1989. Spectral methods of characterizing protein conformation and conformational changes. *In* Protein Structure (T.E. Creighton, ed.) pp. 251-285. IRL Press, Oxford.

Strickland, E.H. 1974. Aromatic contributions to circular dichroism spectra of proteins. *C.R.C. Crit. Rev. Biochem.* 2:113-175.

Swords, N.A. and Wallace, B.A 1993. Circular dichroism analyses of membrane proteins: Examination of environmental effects on bacteriorhodopsin spectra. *Biochem. J.* 289:215-219.

Wingfield, P., Graber, P., Moonen, P., Craig, S., and Pain, R.H. 1988. The conformation and stability of recombinant-derived granulocyte-macrophage colony stimulating factors. *Eur. J. Biochem.* 173:65-72.

Yang, J.T., Wu, C.-S. and Martinez, H.M. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol.*130:208-269.

Key References

Bayley, 1980. See above

A good introduction to the principles and practice of circular dichroism.

Johnson, 1990. See above.

A good account of the practice of far-UV CD spectroscopy and a description and assessment of the determination of protein secondary structure content.

Sears, D.W. and Beychok, S. 1973. Circular dichroism. *In* Physical Principles and Techniques of Protein Chemistry, Part C (S.J. Leach, ed.) pp. 445-593. Academic Press, New York and London.

A comprehensive review of the theory of CD, together with a useful discussion of near-UV spectra in particular for selected proteins.

Strickland, 1974. See above

An older but still excellent account of the practice, spectroscopic basis, and interpretation of the near UV CD spectroscopy of proteins and model compounds.

Yang et al. 1986. See above.

A good account of the practice of far-UV CD spectroscopy and a description and assessment of the determination of protein secondary structure content.

Contributed by Roger Pain
Jožef Stefan Institute
Ljubljana, Slovenia

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Determining the Fluorescence Spectrum of a Protein

Fluorescence spectra provide a sensitive means of characterizing proteins and their conformation. The spectrum is determined chiefly by the polarity of the environment of the tryptophan and tyrosine residues and by their specific interactions. The diminution of intensity of the fluorescence by added "quenchers" provides a measure of the degree to which these residues are buried in the protein structure or exposed to solvent.

One of the main values of fluorescence as a technique for probing protein conformation is that it is highly sensitive and very economical of material. This means, however, that small traces of fluorescent impurities in the solvent or on the cell are readily detected and, if care is not taken, can lead to misinterpretation of the spectra. The essential aim in using this technique, therefore, must be to obtain a fluorescence emission spectrum for the recombinant protein that is guaranteed free from all too easily generated artifacts.

STRATEGIC PLANNING

The following points are essential to obtaining reproducible spectra for reliable comparison between samples.

Fluorescence Spectrometer

Fluorescence spectra are measured with a fluorescence spectrometer (spectrofluorimeter), of which there are a variety of models on the market. For protein characterization, a xenon light source and optics that ensure low levels of stray light are essential. Spectrometers with dual holographic gratings in the excitation monochromator are preferred. There are two main types of fluorescence detection available: analog and photon counting. The latter counts individual photons, which enables high sensitivity and stability, but limits the range of intensities over which a linear response is obtained. Analog spectrometers provide high-precision measurements over a wide range of intensities. Stability is generally high, but is improved by the presence of a reference photomultiplier. In general, the photon-counting spectrometer is for more specialized use and the analog type of instrument is more widely used for the kind of work described in this unit.

A thermostatted cell holder—either one connected to a good circulating thermostat bath or one of the Peltier controlled type—is essential. The latter controls the temperature by electrical means, the advantages being that no water needs to be brought into the spectrometer and, secondly, that the temperature can be raised and lowered very quickly. The spectrometer should incorporate shutters that protect the sample and the photomultiplier from radiation when scanning is not in progress. A magnetic stirrer incorporated into the cell holder is useful to facilitate mixing and to minimize artifacts due to photochemical degradation. Careful attention must be paid to the manufacturer's instructions in order to achieve the optimum settings for measurement.

Calibration of Spectrometer

In order to ensure reproducibility, the wavelength setting should be correct within ±0.5 nm. This will generally not be a problem, but it is just as well to check, most readily with a mercury lamp inserted into the cell compartment. The procedure described by Lakowicz (1983) is recommended. Calibration of intensity can be more difficult, and it is always more reliable to scan samples in parallel with a solution of a known, stable compound, be it the protein in question or a model compound such as *N*-acetyl-L-tryptophanamide.

Contributed by Roger H. Pain
Current Protocols in Protein Science (1996) 7.7.1-7.7.19
Copyright © 1996 by John Wiley & Sons, Inc.

BRIGHAM YOUNG UNIVERSITY LIBRARY

Cells

Choice of cells

The fluorescence cuvettes best used for reliable characterization of proteins are either the standard 10 × 10–mm fluorescence quartz cuvettes (with all four faces polished), taking 2 to 3 ml of solution, or semimicro cuvettes, taking ~0.7 ml. In order to ensure reproducible positioning of the cuvette, it should be placed in the cell holder with the inscription facing the incident light beam.

Preparation and handling of cells

Cuvettes must be thoroughly and reproducibly cleaned by soaking in a proprietary cuvette-cleaning fluid such as Hellmanex II (Hellma), an alkaline liquid concentrate, or RBS-35 Detergent Concentrate (Pierce), following the manufacturer's instructions. The manufacturers of Hellma cuvettes caution against the use of ultrasonics as a means of boosting cleaning efficiency. Fifty percent nitric acid is also effective, but requires greater care and a fume hood. Detergents frequently contain fluorescent material and, if used, should be removed from cuvettes with 2 M HCl followed by water. After the final rinse with copious amounts of high-quality distilled water, the cells are best dried using suction through a Pasteur pipet protected with a small length of plastic tubing. This is safer than using acetone or alcohol, which can leach out fats from fingers and deposit them on the cuvette. The faces of cuvettes must not be touched after cleaning. Filling, emptying, and rinsing between measurements of spectra must be done using pipets, to avoid spillage onto the outer faces. A Pasteur pipet protected with fine plastic tubing is satisfactory. The baseline scans, with buffer in the cell, must be examined as a constant check on the reproducibility of the state of the cell.

The cleanliness of the cuvette can be checked by filling with distilled water and scanning the appropriate wavelength region of the spectrum. If there is no contamination by fluorescent material, the spectrum should be flat, save for the Raman band (see Fig. 7.7.1) on the high-wavelength side of the excitation wavelength, λ_{max}.

Buffer Solutions

Buffers should have low absorbance and low fluorescence in the regions of excitation and emission. Absorbance will usually be expected to be <0.1 and fluorescence close to zero. This will normally be the case for standard buffers made from analytical-grade reagents or materials of equivalent purity but they should, nevertheless, be routinely checked for fluorescence. If a fluorescent component is added to the solution—e.g., as a ligand—it should be checked that the fluorescence observed arises solely from that component. The actual buffer solution used to dissolve or to dialyze the protein should be used for the fluorescence blank. Plastic containers (and stirring bars) may contribute fluorescent agents; if these are used, appropriate blanks should be carefully monitored for fluorescence.

Clarification of Solutions

Protein solutions must be freed from the turbidity resulting from dust and aggregated protein. Many proteins can be filtered (preferably as a stock solution) using a 0.22-μm filter—ideally of the grade special for proteins such as Durapore (Millipore)—fitted to a syringe. Alternatively, if the protein is adsorbed by filters, the solution can be clarified by centrifugation. A microcentrifuge operated at maximum speed for 10 min at 4°C provides reliable clarification for the majority of proteins that do not sediment appreciably under these conditions. Buffer solutions are filtered on a larger scale, using a 0.22-μm filter.

It is essential to monitor clarification if artifacts are to be avoided. There are three indicators of the completeness of clarification, each available from the standard procedures that should be adopted for spectroscopic characterization.

PCU IASST / CENTRAL UNIVERSITY LIBRARY

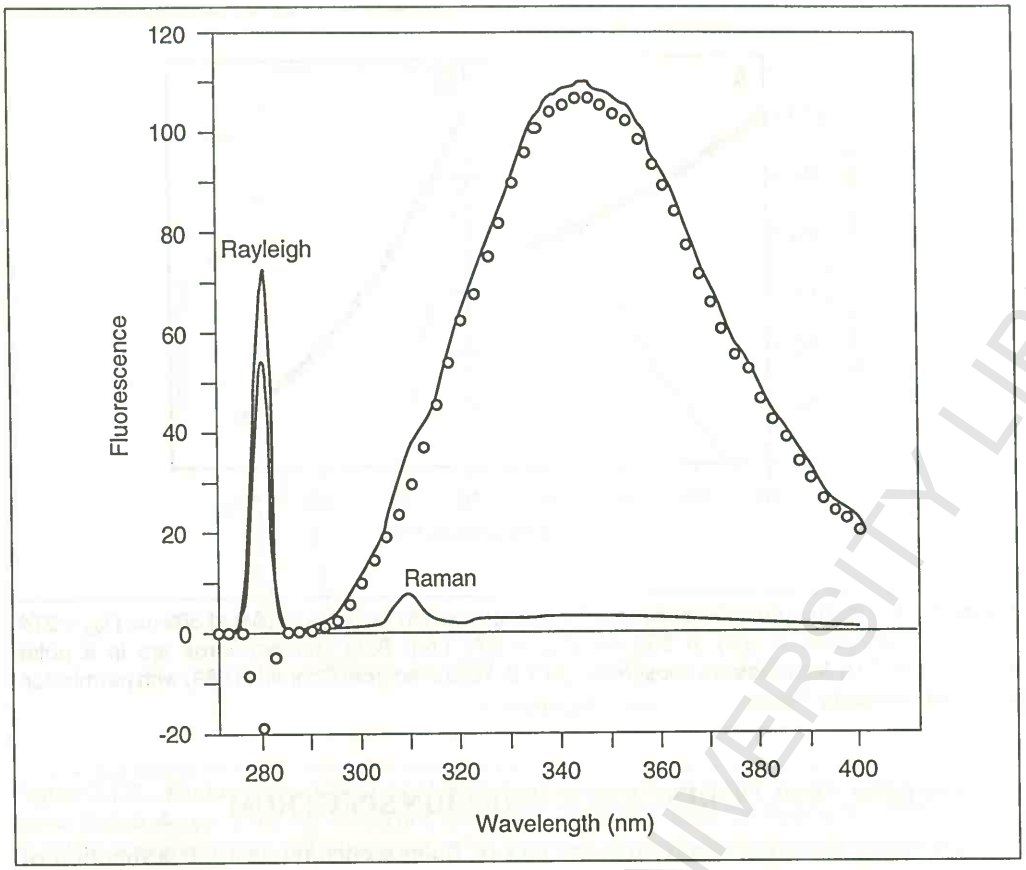


Figure 7.7.1 Fluorescence spectrum of hen egg white lysozyme (EWL). The Rayleigh and Raman bands are seen in the scans for the solvent baseline and protein solutions (solid lines). Circles represent the spectrum of EWL with baseline subtracted. Parameters: EWL $A_{280} = 0.05$; $\lambda_{ex} = 280$ nm; excitation and emission bandwidths, 2.5 nm; scan rate, 100 nm/min; five scans accumulated. Spectrum was measured using a Perkin Elmer LS50B fluorescence spectrometer.

1. Determination of protein concentration (*UNIT 7.2*) requires an absorbance spectrum to be recorded on a good quality spectrophotometer from 240 to 350 nm. Aromatic amino acid residues do not absorb above 320 nm, so the spectrum between 320 and 350 nm should be only marginally above baseline. The presence of turbidity will result in finite attenuance in this region that increases toward lower wavelengths.

2. From the above absorbance spectrum, the ratio A_{max}/A_{min} is measured, where A_{min} is the absorbance at the trough in the region 240 to 250 nm. This ratio will be specific for a given protein and should be measured for a sample that has been carefully and reproducibly clarified. Due to the strong wavelength dependence of scattering, the ratio is very sensitive, decreasing sharply with turbidity, and can be used as a semiquantitative check on the clarity of further solutions of the protein.

3. The scan of a fluorescence emission spectrum should routinely include the region of the Rayleigh peak (see Fig. 7.7.1). This provides a built-in indicator of turbidity of the buffer or protein solutions. The height of the peak should not differ greatly from that obtained for the solvent from the equivalent buffer blank scan; it will increase markedly in height if there is dust or aggregate in the solution.

BCU LSI - CENTRAL UNIVERSITY LIBRARY

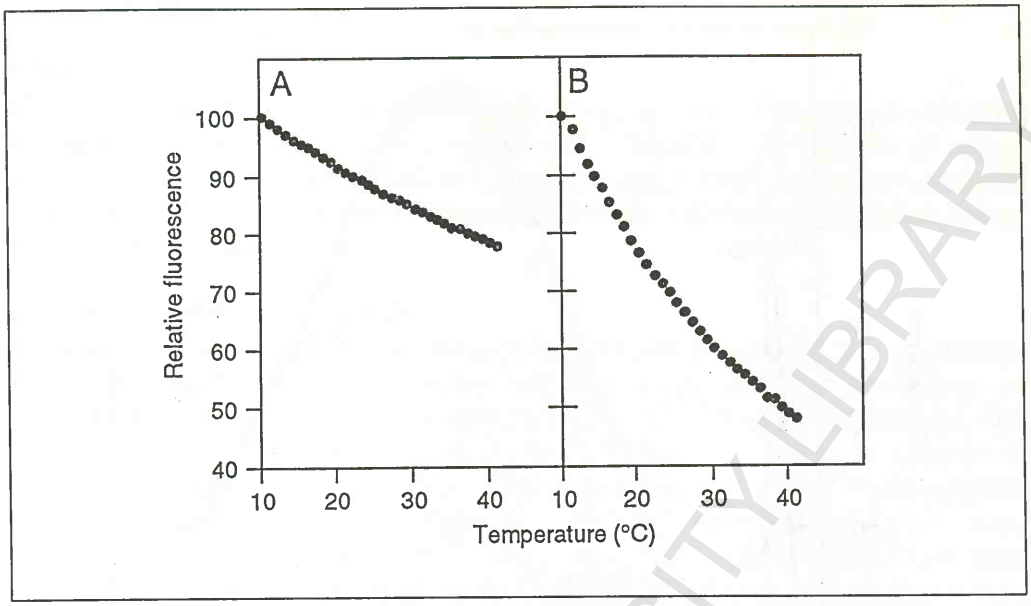


Figure 7.7.2 Temperature dependence of fluorescence. (**A**) Tyrosine (6 μM) at 303 nm (λ_{ex} = 274 nm). (**B**) Tryptophan (1 μM) at 355 nm (λ_{ex} = 278 nm). Both chromophores are in a polar environment, 0.01 M potassium phosphate, pH 7.0. Reprinted from Schmid (1989) with permission of Oxford University Press.

BASIC PROTOCOL 1

RECORDING A FLUORESCENCE EMISSION SPECTRUM

A fluorescence emission spectrum is the plot of fluorescence intensity as a function of wavelength. Fluorescence emission takes place from the wavelength of excitation (λ_{ex}) upwards. The range of the spectrum recorded is described below (step 5).

Materials

Calibrated fluorescence spectrometer, buffer solution, and cleaned cuvettes (see Strategic Planning)

Clarified protein solution of known concentration (see Strategic Planning for clarification methods and *UNIT 7.2* for concentration determination) with absorbance usually in the range of 0.05 to 0.1 at the wavelength to be used for excitation

Prepare spectrometer

1. Switch on the spectrometer and allow it to warm up for the time specified in the manufacturer's instructions (usually ~30 min). Regulate the thermostat control to the desired temperature (normally 25°C unless the stability of the protein dictates otherwise).

   The intensity of fluorescence is highly temperature-dependent (see, e.g., Fig. 7.7.2) and it is important that the temperature be accurately controlled and measured in the cuvette with a calibrated electric thermometer. It is useful for future reference to note the time taken for the sample to reach the operating value.

2. Set the excitation wavelength (λ_{ex}).

   Setting λ_{ex} to the λ_{max} of the absorbance spectrum (usually 280 nm) will excite both tyrosine and tryptophan residues. Alternatively, λ_{ex} may be set to 295 nm in order to excite only tryptophan (see annotation to step 3).

3. Set the excitation bandwidth.

   The term bandwidth is described and defined in UNIT 7.6, Basic Protocol step 2b. The excitation bandwidth (in nm) is determined by the excitation slit width (in mm). It will

PCU IASI CENTRAL UNIVERSITY LIBRARY

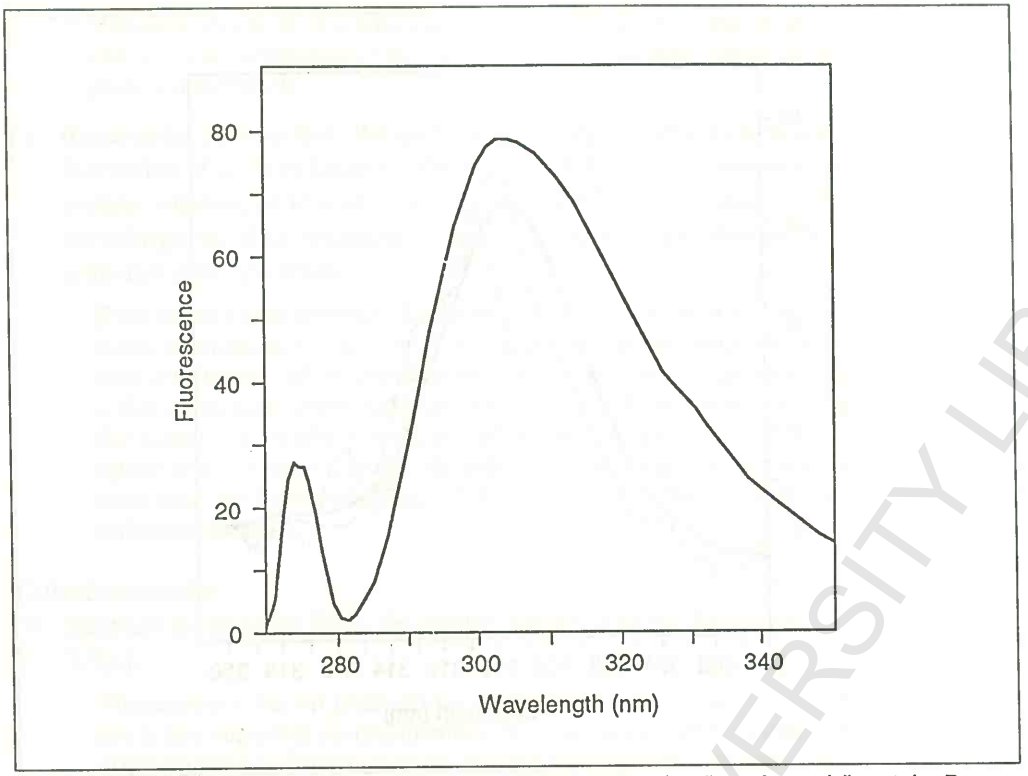


Figure 7.7.3 Fluorescence spectrum of ribonuclease (RNase), a "tyrosine only" protein. Parameters: Rnase A_{275} = 0.05; λ_{ex} = 275 nm; one scan, smoothed; other settings as for Figure 7.7.1. Baseline was measured with the same settings and subtracted. Measurements made using a Perkin-Elmer LS50B.

determine the intensity of the excitation and the amount of scattered light from the Rayleigh peak that will be "seen" in the emission spectrum. It is important in certain cases in determining which chromophore is excited. It is particularly important, for example, to limit the bandwidth when measuring the contribution of tyrosine in the presence of tryptophan. The emission maximum of tyrosine at 303 nm (see Fig. 7.7.3) is close to the required excitation wavelength of 295 nm. The envelope of the Rayleigh peak must not, in this case, significantly overlap the emission band of tyrosine. Settings of 2.5 to 5 nm for $\lambda_{ex} \approx 280$ nm and ≤ 2.5 nm for $\lambda_{ex} \approx 295$ nm are usual.

4. Set the emission bandwidth.

Since protein emission spectra are generally rather broad, larger emission bandwidths can usually be tolerated. Only where it is important to resolve tyrosine fluorescence from tryptophan fluorescence and the Rayleigh scattering peak is it necessary to minimize the bandwidth. In cases where only small amounts of material are available, it is necessary to strike a balance between resolution and light intensity in order to obtain the best possible signal-to-noise ratio. Settings of 2.5 to 10 nm are normal.

5. Set the range of emission wavelengths to be scanned for the emission spectrum.

The lower wavelength should be ~10 nm below the excitation wavelength in order to record the Rayleigh peak (see discussion of Clarification of Solutions in Strategic Planning). The higher end of the range should be chosen to include as much of the spectrum as possible, so as to allow integration of the peak if necessary. Values of 400 and 340 nm are suitable for proteins with and without tryptophan, respectively. Hence, for example, a typical emission wavelength range for a Trp-containing protein excited at 295 nm would be 285 to 400 nm.

BOSTON UNIVERSITY LIBRARY

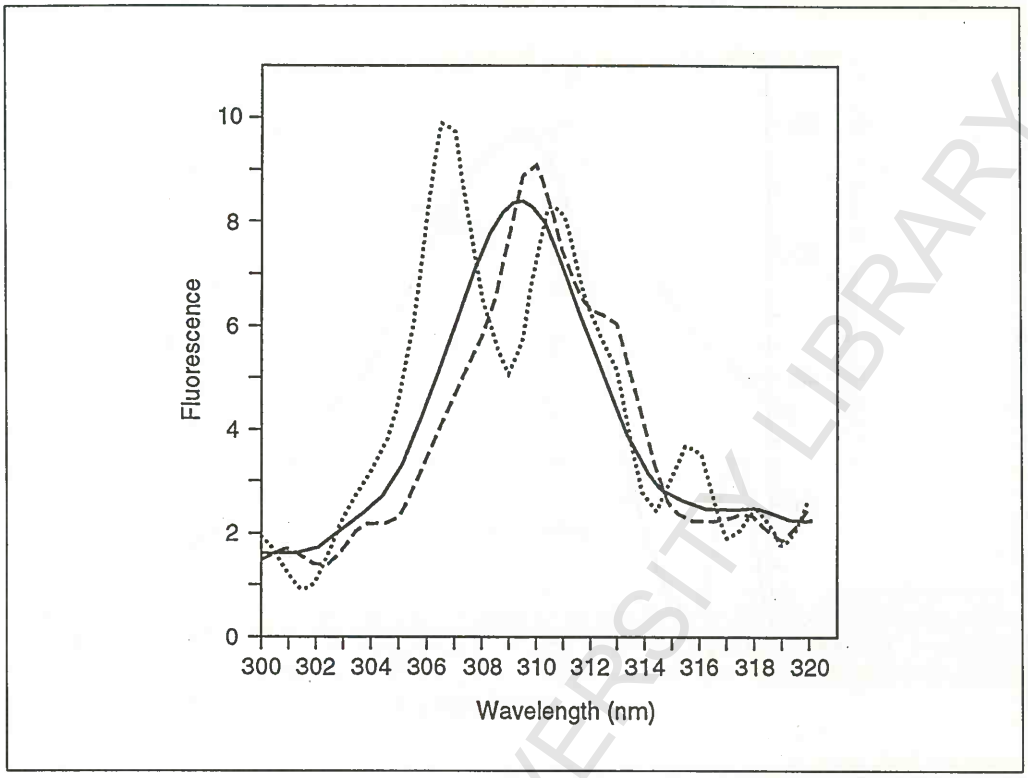


Figure 7.7.4 The Raman band measured under different scanning conditions. The appropriate section of the baseline in Figure 7.7.1 is shown for: a single scan at 1500 nm/min (dotted line); ten scans at 1500 nm/min (dashed line); and ten scans at 100 nm/min (solid line). The times required for scanning the complete baselines for these three measurements were 0.087 min, 0.87 min, and 13 min, respectively. Measurements made using a Perkin-Elmer LS50B.

6. Set the scan rate.

   The scanning rate for the emission wavelength has to be slow enough to allow the instrument to respond to changes in intensity with wavelength. Too fast a scan, even with a number of accumulations, results in an apparent spectrum that can differ markedly from the true spectrum (Fig. 7.7.4). When, as is usually the case with proteins, the emission band is rather broad (Fig. 7.7.1), a scan rate setting in the region of 100 nm/min will usually be appropriate.

Scan buffer blank and protein sample

7. Fill a cleaned cuvette with a filtered sample from the actual batch of buffer used to dissolve or dialyze the protein, then place the cuvette in the cell holder of the spectrometer in a reproducible orientation (see discussion of Cells in Strategic Planning). Scan this buffer blank using the same instrument settings as are appropriate for the sample, store the spectrum, and check for any unexpected fluorescence bands.

   If the settings appropriate for the protein sample cannot be estimated in advance, the sample may be run before the buffer blank, although special care will have to be taken to remove any protein adsorbed to the inner faces of the cuvette.

   Temperature equilibration is not usually important for the buffer, except in cases where it contains fluorescent components.

   Cell holders are susceptible to corrosion by salt solutions. Cells should therefore be filled outside the spectrometer.

   Apart from the Raman band, a buffer solution should give an almost flat baseline of low intensity (see Fig. 7.7.1). At the usual working concentrations of protein (A_{280} ≥0.05), buffer fluorescence constitutes only a minor contribution to the fluorescence spectrum of the protein and is accounted for by the appropriate baseline subtraction.

CU IASI CENTRAL UNIVERSITY LIBRARY

It is advisable to repeat baseline scans at intervals during a series of experiments to check on contamination of the outer faces of the cuvette, and of the inner faces due to protein adsorption.

8. Remove the cuvette from the spectrometer, empty with a protected Pasteur pipet (see discussion of Cells in Strategic Planning) and refill in the same way with the clarified protein solution of known concentration with an absorbance of 0.05 to 0.1 at the wavelength used for excitation. Allow to come to temperature in the cell holder, then scan and store spectrum.

   If the protein concentration has to be low, and/or the intensity of fluorescence is low, it may be necessary to accumulate a larger number of scans. Proteins differ widely in their sensitivity to photodegradation, but exposure to UV radiation may lead to damage to those molecules in the light path, with subsequent reduction in fluorescence intensity. Excitation slit widths should, therefore, be as narrow as is consistent with good signal-to-noise ratio. Stirring the solution in the cuvette will bring undamaged molecules into the light path (Fig. 7.7.8) and further reduce significant artifacts due to radiation damage.

Calculate results

9. Subtract the baseline from the sample spectrum using the instrument software (Fig. 7.7.1).

   Fluorescence has no units. Hence, if the fluorescence intensities of different samples are to be compared, careful attention must be paid to reproducible concentrations and instrumental conditions. Instrument response may vary with time, so must be controlled by the use of stable standards.

10. Correct the apparent fluorescence intensities for the inner filter effect.

    A protein solution will absorb a certain proportion of the radiation before it reaches the volume of solution whose fluorescence is viewed by the optics of the instrument (see Fig 7.7.5). For a solution of absorbance 0.1, the intensity at the center of the cell is about 10% less than that of the incident radiation, with a corresponding decrease in the observed fluorescence. Similarly, the fluorescence intensity may be reduced by absorption in the exit path. This is known as the inner filter effect. In practice, it means that the observed fluorescence intensity is directly proportional to protein concentration only at low concentrations, when absorbance is <0.1 at λ_{ex}. The apparent fluorescence intensity (F_{app}) can be corrected by the following approximation:

$$F_{corr} = F_{app} \text{ antilog}\{(A_{ex} + A_{em})/2\}$$

    where A_{ex} and A_{em} are the solution absorbances at the excitation and emission wavelengths, respectively.

    The effect is much more marked if a third component that absorbs in the wavelength range of the protein absorption or fluorescence is added to the solution (see Basic Protocol 2).

11. Determine the value of the fluorescence intensity at the wavelength of maximum intensity (I_{max} at λ_{max}), using, where available, the instrument software.

    It is possible to obtain corrected spectra, which take into account the variation of intensity of the light source as well as variation of the detector response with wavelength. The procedures for calculating these are not simple (see Lakowicz, 1983), and for the purposes of characterization described here the uncorrected spectra are normally used. Because of the above variations, however, it is important when comparing spectra to do so with spectra obtained using the same instrumental settings and, preferably, the same instrument.

    In order to compare complete spectra from different samples of a protein, they should be normalized by dividing the value of fluorescence intensity by the protein concentration (i.e., I_{em}/c).

BOLTON / CENTRAL UNIVERSITY LIBRARY

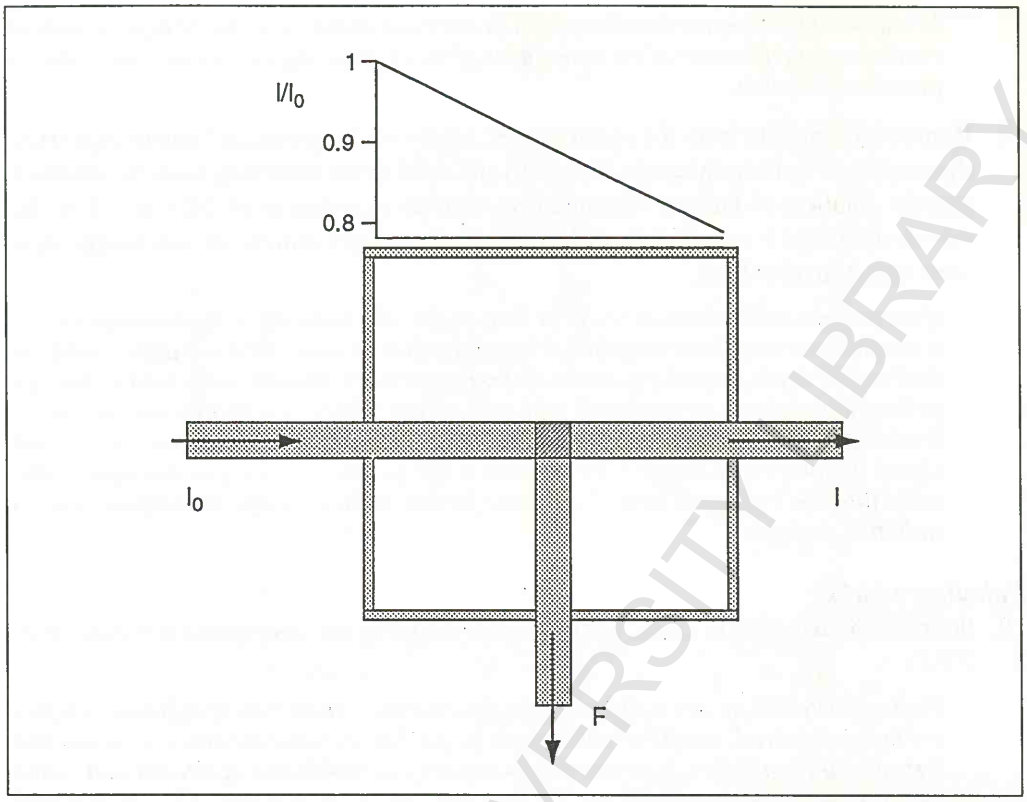


Figure 7.7.5 The inner filter effect. A cuvette (10 × 10–mm) is represented in plan view, with the collimated incident beam from the monochromator having intensity I_0. As a result of absorption by the protein solution, the intensity through the cuvette will decrease steadily, emerging with intensity I. The values are illustrated for a solution having absorbance at the excitation wavelength of 0.1. The optics of the fluorescence detector are focused so that only the fluorescence originating from the volume depicted by the heavily shaded square is "seen" by the photomultiplier. Thus the observed fluorescence intensity will be less than that expected from the fluorescence of the protein at infinite dilution. The fluorescence passes through the protein solution on its way to the detector and will be further decreased in intensity if the solution absorbs at the wavelengths of the emitted radiation.

BASIC PROTOCOL 2

DETERMINATION OF FLUORESCENCE QUENCHING

Fluorescence quenching provides a means of probing the accessibility of tryptophan residues to small molecules and thus gives information about their structural environment. The technique involves quantifying the decrease in protein fluorescence intensity in the presence of increasing concentrations of quencher, followed by analysis of the data to give details of the interaction of the quencher with the tryptophan residue.

Materials

Protein solution in buffer, A_{280} = 0.05 to 0.1

Quenchers: these are most conveniently made up as concentrated stock solutions; examples of frequently used quenchers include:

5 M NaI or 2.5 M KI solution containing 1 mM $Na_2S_2O_3$ to prevent I_3^- formation

5 M CsCl (optical grade, Aldrich)

8 M acrylamide (Electran grade from BDH; ε_{295} = 0.236 liter/mol/cm)

2.5 M succinimide (recrystallized from ethanol with activated charcoal treatment; ε_{295} = 0.03 liter/mol/cm)

Additional reagents and equipment for recording a fluorescence spectrum (see Basic Protocol 1)

FCU IASL / CENTRAL UNIVERSITY LIBRARY

1. Scan the fluorescence spectrum of a measured volume of the protein in the required buffer (see Basic Protocol 1).

2. Add a known small volume of a stock solution of quencher, mix thoroughly and again scan the spectrum.

   Alternatively, for economy of time, measure the intensity of fluorescence at just two wavelengths—315 and 350 nm—allowing a sufficient time of accumulation to achieve a reproducible average reading.

   The actual amount of quencher added is a matter of judgment, depending on the degree of quenching found. It is important to obtain at least ten measurements of fluorescence intensity spread evenly between zero and full quenching (see Fig. 7.7.6).

   Quenchers should be carefully purified. When iodide is used, it should be in the presence of 1 mM $Na_2S_2O_3$ to avoid the formation of I_3^-, which can react with tyrosine.

   The tryptophan residues in a given protein are likely to be in different environments and, therefore, to exhibit different values of λ_{max}. Measurement of the degrees of quenching at the red and blue ends of the emission spectrum will determine whether this is so and will also add further detail to the overall fingerprint of the protein (see discussion of Quenching in Support Protocol).

3. Correct each reading of the fluorescence intensity for dilution and for the absorbance of the quencher (see Basic Protocol 1).

   See also Eftink and Ghiron (1981).

4. Plot the resulting fluorescence intensities as a function of concentration of quencher according to the Stern-Volmer equation (see Fig. 7.7.6A):

   $$F_0/F = 1 + K_D[Q]$$

   where F_0 and F are the fluorescence intensities in the absence and presence, respectively, of a given concentration of quencher $[Q]$, and K_D is the Stern-Volmer quenching constant.

   K_D is determined by the bimolecular quenching constant (k_q) and by the fluorescence lifetime of the fluorophore (τ_0), in the absence of quencher ($K_D = k_q\tau_0$). The slope of a plot of F_0/F against $[Q]$ gives K_D directly, in units of liters/mol, with an intercept of unity. This plot is expected to be linear for a single fluorophore in a single environment. For a more complex system the plot will be non-linear (see Fig. 7.4.6A) and the data should be plotted according to a modified Stern-Volmer equation:

   $$F_0/\Delta F = 1/f_a K_D[Q] + 1/f_a$$

   where f_a is the fraction of the initial fluorescence accessible to the quencher,

   $$f_a = F_{0a}/(F_{0a} + F_{0b})$$

   and K_D is the Stern-Volmer quenching constant for that fraction of the fluorescence. A plot of $F_0/\Delta F$ against $1/[Q]$ gives an intercept of $1/f_a$ and a slope of $1/f_aK_D$ (see Fig. 7.7.6B).

B. C. U. „M. EMINESCU" IAŞI
CHILIE

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

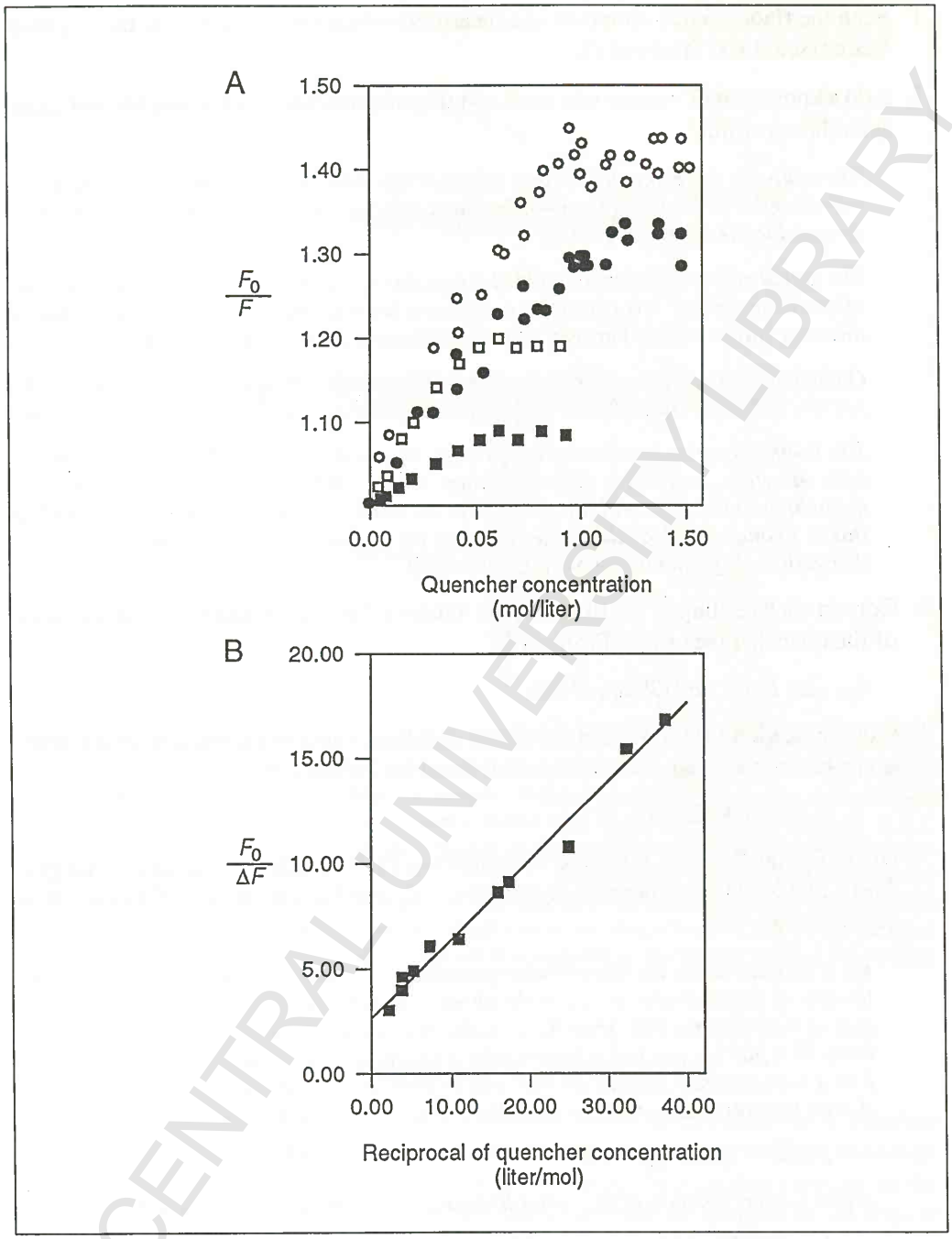


Figure 7.7.6 Quenching of fluorescence of 3-phosphoglycerate kinase (PGK). (**A**) Fluorescence emission intensities, expressed as a fraction of the unquenched fluorescence, are plotted according to the Stern-Volmer equation (see Basic Protocol 2, step 4) at 316 nm (filled symbols) and 350 nm (open symbols). Quenchers used were succinimide (circles; recrystallized from ethanol and treated with activated charcoal, 2.5 M stock solution) and potassium iodide (squares; 2.5 M stock solution with 1 mM $Na_2S_2O_3$). Fluorescence intensities were adjusted for dilution and for quencher absorbance using a molar extinction coefficient $\varepsilon_{295} = 0.03$ liter/mol/cm for succinimide. The greater level of quenching at 350 nm reflects the fact that the more exposed of the two tryptophan residues has a more red-shifted fluorescence maximum. (**B**) Modified Stern-Volmer plot for the quenching of PGK fluorescence by succinimide. The data in panel A for quenching at 350 nm is plotted according to the modified Stern-Volmer equation (see Basic Protocol 2, step 4 annotation). The intercept on the *y*-axis is 2.36 ± 0.2. The results in panels A and B are characteristic of a protein where part of the fluorescence is inaccessible to a given quencher, in this case, iodide and succinimide. Measurements made using a Perkin-Elmer MPF3-L. Reprinted from Varley (1991) with permission of Elsevier Science.

FCU IAS CENTRAL UNIVERSITY LIBRARY

BASIC THEORY AND INTERPRETATION OF FLUORESCENCE SPECTRA

When radiation is incident on a solution, most of it is scattered without change in wavelength. Thus, exciting radiation at 280 nm will be scattered and a fraction of it will be detected by the spectrofluorimeter as a peak at the same wavelength. The width of this **Rayleigh** scattering peak (Fig. 7.7.1) will depend on instrumental settings, mainly those of the excitation and emission slit widths, which determine the range of wavelengths observed. The **bandwidth** is the width, in nm, of the distribution of intensity measured at half the height of the maximum intensity.

A further low-intensity peak, known as the **Raman band**, is also observed in spectra (see Fig. 7.7.1). This is a low-intensity band of scattered radiation whose distance from the excitation band is a measure of the vibrational energy of the H–O bond in solvent water. At λ_{ex} = 280 nm, the Raman band for water occurs at 311 nm; in general, it occurs at $1/\{(1/\lambda_{ex}) - (3.6 \times 10^{-4})\}$, where λ_{ex} is the excitation wavelength in nm. The intensity and resolution of the Raman band provides a useful empirical check on the performance of the spectrofluorimeter. A decrease in the signal-to-noise ratio usually indicates a deterioration of the lamp.

Chromophores, such as the sidechains of tryptophan and tyrosine, will absorb some of the radiation in ~10^{-15} sec and become excited from the ground electronic state (S_0) to higher vibrational levels in a higher energy electronic state (S_1; see Fig. 7.7.7). The energy in the vibrational levels is rapidly (~10^{-12} sec) lost to the surroundings as thermal energy, so that the excited chromophore is very largely in its ground vibrational level. Return to the ground electronic state can result in the emission of radiation of lower energy and hence longer wavelength (Figs. 7.7.1 and 7.7.3), or the energy can be dissipated in a

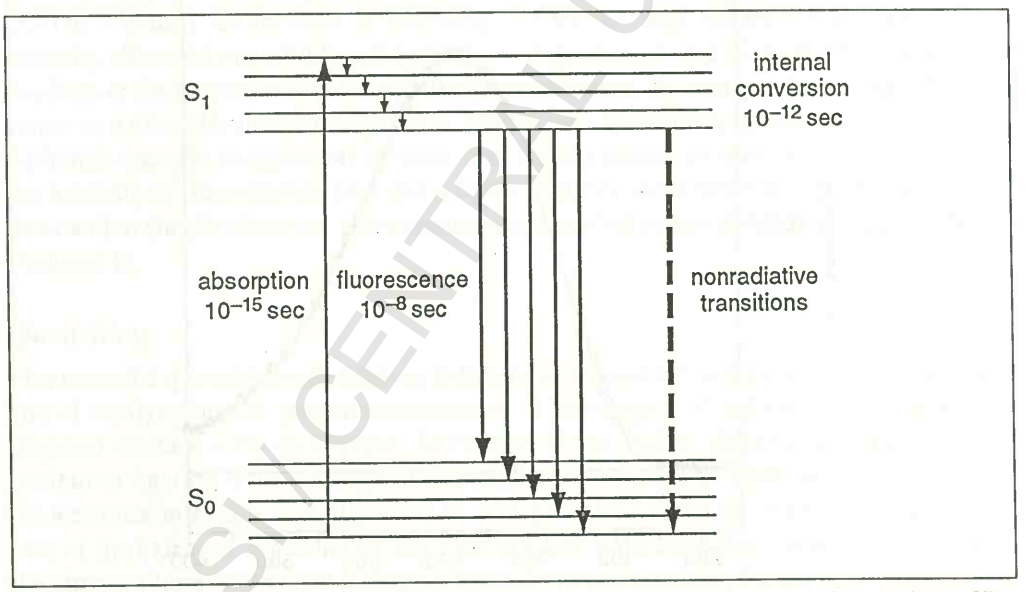


Figure 7.7.7 The processes of absorption and emission of radiation by a chromophore. The electronic structure of a chromophore is in the ground state (S_0) under normal conditions. Absorption of radiation, a rapid process, leads to excitation into an excited state (S_1). Within each electronic state are a number of vibrational states, represented by horizontal lines. Under normal thermal conditions, only the lower one or two vibrational levels are occupied. Thus, although excitation takes place into a range of levels in the excited electronic state, emission, in the form of fluorescence, takes place from the lowest vibrational level. The lengths of the vertical lines representing absorption and emission illustrate the relative energies of the radiation in each process and illustrate why fluorescence occurs almost always at longer wavelengths than absorption. The vertical broken line represents the dissipation of absorbed energy by means other than radiation.

PRINCE'S CENTRAL UNIVERSITY LIBRARY

number of ways that are nonradiative (Fig. 7.7.7), including quenching (see discussion of Quenching, below). The ratio of the rate constants for these two alternative processes defines the quantum yield or fluorescence efficiency, which is almost always less than unity.

The chromophores tyrosine and tryptophan absorb radiation in the 280 nm region, re-emitting a proportion of it as fluorescence. The quantum yields for tryptophan, tyrosine and phenylalanine are 0.2, 0.1, and 0.04, respectively. Coupled with their relative absorption coefficients (*UNIT 7.2*), it can be seen why most protein fluorescence spectra are dominated by tryptophan.

Quantitation of the fluorescent properties of a protein can provide a number of rather specific parameters that allow assessment of the correct folding, besides giving additional structural information.

Tyrosine Fluorescence Is Frequently Not Seen in the Presence of Tryptophan

The fluorescence intensity of free tyrosine is about one-fifth of that of tryptophan, but in proteins it is usually much weaker still. This is due to a combination of interactions—such as hydrogen bonding to peptide carbonyl groups, the presence of neighboring charged groups, and nonradiative energy transfer to tryptophan or disulfide bonds—all of which can result in tyrosine fluorescence being diminished or quenched. The degree of quenching varies from protein to protein and can be observed by paying careful attention to bandwidths (see Basic Protocol 1). In Figure 7.7.8, where the peak intensity at 335 nm of

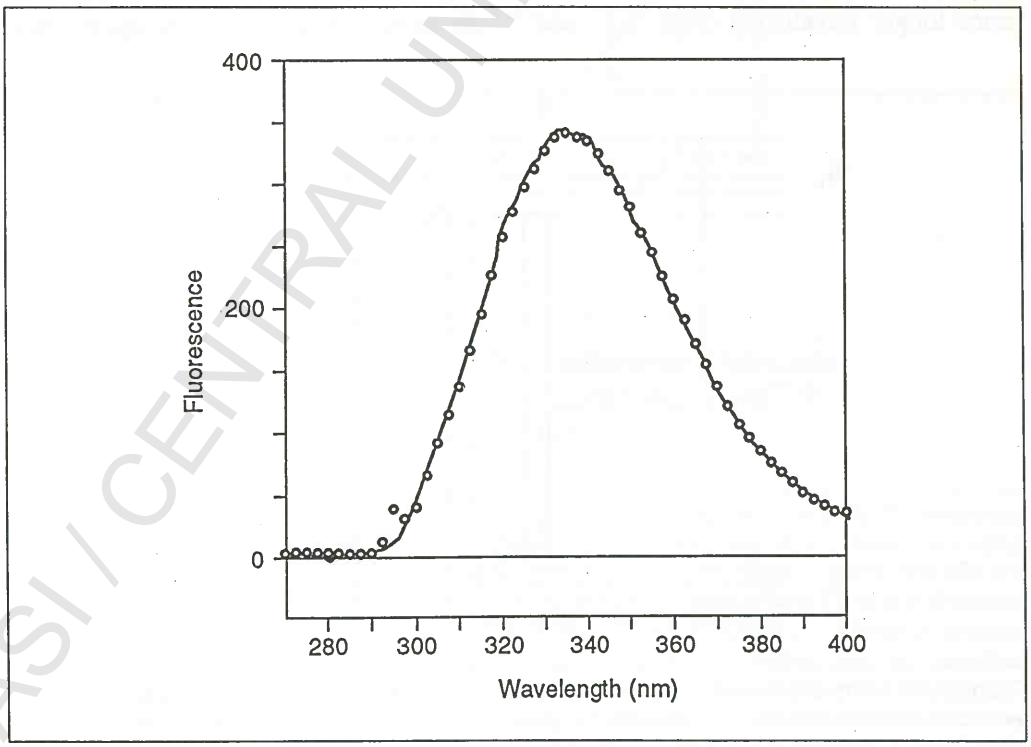


Figure 7.7.8 Quenching of the tyrosine fluorescence of chymotrypsinogen A. The solid line represents the emission spectrum excited at 280 nm. The circles represent the corresponding spectrum obtained with λ_{ex} = 295 nm. Because the latter spectrum is considerably less intense than the former (as a result of the differences in absorbance at 280 nm and 295 nm), the spectra were normalized using the ratio of their fluorescence intensities at 335 nm. Parameters: chymotrypsinogen A_{280} = 0.1; excitation and emission bandwidths, 2.5 nm; scan rate, 100 nm/min, two scans accumulated for λ_{ex} = 280 nm and four scans for λ_{ex} = 295 nm. Fluorescence was measured using a Perkin Elmer LS50B.

FCU IASI / CENTRAL UNIVERSITY LIBRARY

fluorescence excited at 295 nm (i.e., tyrosine is not excited) has been normalized to that excited at 280 nm (where both tryptophan and tyrosine are excited), it is evident that there is no separate contribution of tyrosine to the fluorescence excited at the latter wavelength. Essentially all the tyrosine fluorescence is quenched (see Lakowicz, 1983). The unfolded protein shows a separate shoulder resulting from unquenched tyrosine fluorescence.

λ_{max} and I_{max} Reflect the Detailed Environment of Aromatic Residues

The precise wavelength of emission is, in the case of tryptophan, dependent on the polarity of the immediate environment of the chromophore in the protein, and maximum values have been observed over a broad range between 308 nm (highly nonpolar) and 352 nm (fully exposed to solvent). In tumor necrosis factor (TNF), the maximum emission wavelength is 320 nm, compared to 350 nm for the fully unfolded protein (Fig. 7.7.9). The value of 336.5 nm for λ_{max} in acid solution indicates partial unfolding. By comparison, tryptophan residues in native chymotrypsinogen A, where λ_{max} = 332 nm (Fig. 7.7.10), are in a slightly less nonpolar environment than in TNF. The value of λ_{max} therefore provides both a characteristic parameter for a particular protein and structural information concerning the average degree of burial of the tryptophans. The wavelength of tyrosine fluorescence, on the other hand, does not vary with polarity of environment and occurs with a maximum at 303 nm. Phenylalanine does not absorb above 275 nm, so that its weak fluorescence is not normally observed.

The intensity of fluorescence is, for both tyrosine and tryptophan, dependent upon interactions with neighboring groups that may quench the fluorescence emission from the native conformation. Tryptophan and tyrosine fluorescence is quenched, relative to the emission intensity in water, in solutions of denaturant. Depending on the degree of quenching of these chromophores in the folded protein, therefore, denaturation can result in an increase or a decrease in intensity, as well as a shift in λ_{max} to longer wavelengths (350 to 355 nm). In the case of TNF (Fig. 7.7.9), the large reductions in fluorescence intensity observed on unfolding the protein in guanidine or acid show that the tryptophan residues in the native protein are relatively unquenched. By contrast, tryptophan fluorescence in native chymotrypsinogen A is seen to be considerably quenched (Fig. 7.7.10). Although specific assignments of these interactions cannot be derived from the spectra, the intensity of fluorescence provides a further empirical characteristic of a protein. Note that caution should observed when comparing recorded values of λ_{max} and I_{max} (see Basic Protocol 1).

Quenching

Fluorescence quenching provides an indicator of the solvent accessibility of tryptophans and of the dynamics of protein conformation. The presence of certain extrinsic groups in physical contact with an excited chromophore can lead to sharing or transfer of the excitation energy, with consequent reduction in the quantum yield and diminution of fluorescence intensity. The efficiency of this quenching depends on the electronic structure of the extrinsic molecule or ion. The fact that it depends also on its accessibility to the chromophore means that a study of quenching can provide a further and independent probe of tryptophan environment. Iodide ion (I^-) cesium ion (Cs^+), and succinimide are useful quenchers for determining accessibility (a fuller list of quenchers for proteins is provided by Eftink and Ghiron, 1981). The ionic character of I^- and Cs^+ and the relatively large size of succinimide mean that these quenchers cannot readily penetrate the nonpolar interior of the protein, so they can be used to differentiate between buried and solvent-accessible chromophores. Their accessibility to an aromatic residue in the surface of the protein will depend, in addition, on the stereochemistry, chemical nature, and, in particular, charge of the surrounding groups. Quenching can be readily quantitated using the

BOSTON BIOMEDICAL CENTRAL UNIVERSITY LIBRARY

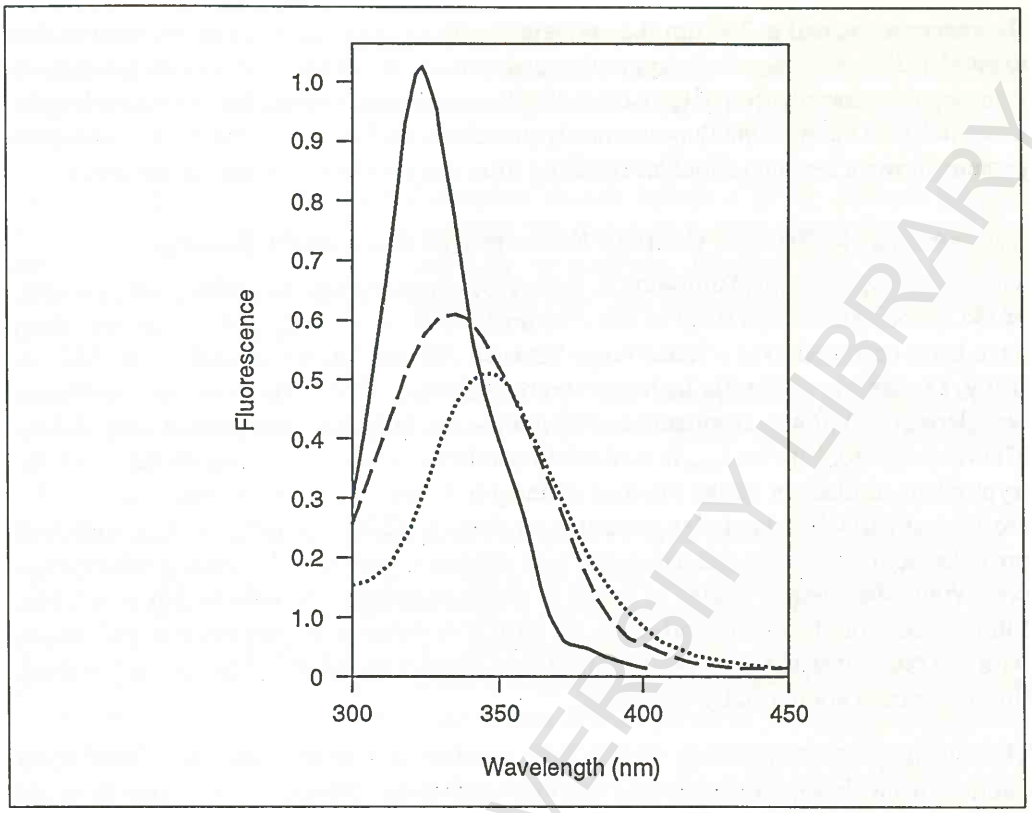


Figure 7.7.9 Quenching of fluorescence of tumor necrosis factor (TNF) upon denaturation. Fluorescence emission spectra are shown for native TNF (solid line), guanidine-unfolded TNF (dotted line), and acid-denatured TNF (dashed line). Parameters: TNF concentration, 30 µg/ml; λ_{ex} = 280 nm; bandwidths ranging from 16 to 24 nm. Measurements made using Perkin-Elmer MPF3 spectrofluorimeter. Reprinted from Hlodan and Pain (1994) with permission of Elsevier Science.

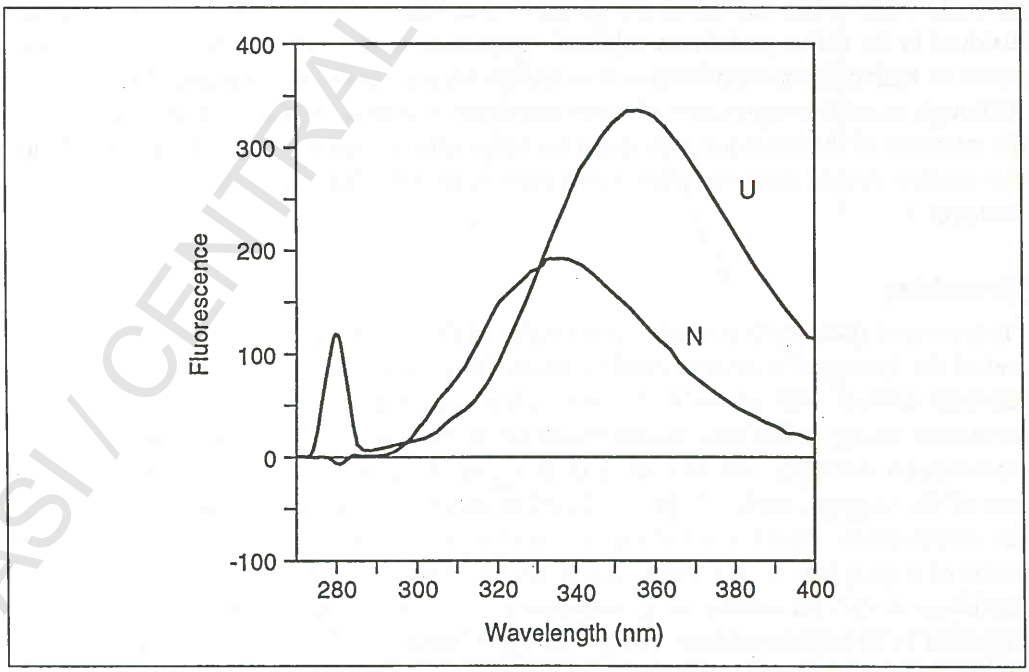


Figure 7.7.10 Dequenching of chymotrypsinogen A on unfolding. Native chymotrypsinogen A (N) was scanned as in Figure 7.7.8. Unfolded chymotrypsinogen A (U) was scanned in 6 M guanidinium chloride using the following parameters: A_{280} = 0.05; λ_{ex} = 280 nm; excitation and emission bandwidths, 2.5 nm; scan rate 100 nm/min; single scan, smoothed. The tryptophan fluorescence is seen to be considerably quenched in the folded protein. Measurements made using a Perkin-Elmer LS50B.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Stern-Volmer equations to give the quenching constant (K_{SV}) and the fraction of the initial fluorescence that is accessible to the quencher (see Basic Protocol 1).

The quenching of fluorescence of phosphoglycerate kinase (PGK) is shown plotted according to the Stern-Volmer equation (see Fig. 7.7.6A). The quenching levels off above 0.9 M succinimide and 0.6 M iodide, respectively, indicating that part of the fluorescence is unquenched by these quenchers. The different degrees of quenching by iodide and succinimide is due to the nature of the neighboring groups affecting access to the tryptophan. The plot of some of the data according to the modified Stern-Volmer equation (Fig. 7.7.6B) shows the fluorescence quenched by succinimide to be 42% of the total. The different degrees of quenching at 316 and 350 nm are due to the two tryptophans having substantially different values of λ_{max}, and hence different degrees of accessibility to solvent. Thus, taking into account the crystal structure of PGK, one of the two tryptophans, W333, is seen to be buried and inaccessible to quencher while the other, W308, is quenched with a quenching constant of $K_{SV} = 0.39$ liters/mol.

Another quencher, acrylamide, can quench the fluorescence not only of exposed but also of buried tryptophan residues based on its ability to diffuse through the protein structure. The efficiency of quenching is, therefore, a reflection of the tightness of packing and the dynamics of the protein. Figure 7.7.11 shows the results of acrylamide quenching of the

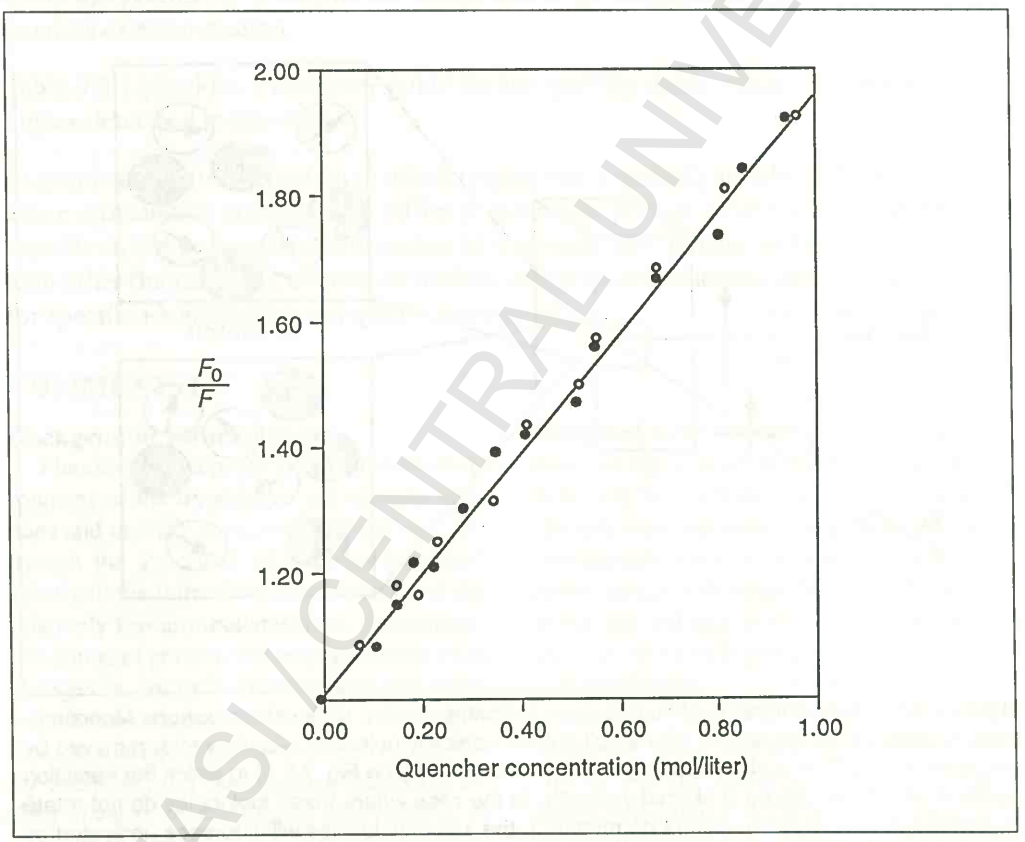


Figure 7.7.11 The quenching of buried tryptophan fluorescence in 3-phosphoglycerate kinase (PGK). In the presence of 1.12 M succinimide (shown to quench the exposed tryptophan fluorescence; see Fig. 7.7.6), fluorescence is further quenched by acrylamide, which is able to gain access to the buried tryptophan. Fluorescence was measured at 316 nm (filled circles) and at 350 nm (open circles) and plotted according to the Stern-Volmer equation (see Basic Protocol 2, step 4). Fluorescence intensities were corrected for dilution and for quencher absorbance using $\varepsilon_{295} = 0.236$ for acrylamide. The linear plot indicates the quenching of a single residue. Measurements made using a Perkin-Elmer MPF3-L. Reprinted from Varley (1991) with permission of Elsevier Science.

BOSPHORUS UNIVERSITY CENTRAL LIBRARY

buried W333 of PGK (in this experiment the more exposed W308 is quenched by succinimide). The linearity of the plot, combined with the coincidence of the quenching measured at 316 and 350 nm (in contrast to Fig. 7.7.6), shows that a single tryptophan only is being quenched by the acrylamide. A quenching constant $K_{SV} = 0.94$ liters/mol is obtained from the slope. In general, bimolecular rate constants for quenching, which reflect more directly the dynamics of the protein, can also be calculated if fluorescence lifetime measurements are available (see discussion of Stern-Volmer equation in Basic Protocol 2; see also Eftink and Ghiron, 1981).

Fluorescence Depolarization

If a polarizer is placed in the excitation beam, the radiation passing through the sample solution will be plane-polarized. This beam will now be absorbed only by those chromophores whose transition dipoles are in the same plane (see Fig. 7.7.12). If the chromophore does not rotate within 10^{-8} sec (see Fig. 7.7.7), the fluorescence will be emitted with the same degree of polarization and will pass through a second polarizer (or analyzer) oriented parallel to the first polarizer. If, on the other hand, the excited chromophores rotate before fluorescence takes place, emission will take place from a smaller number of molecules reoriented with transition dipoles parallel and perpendicular to the exciting radiation

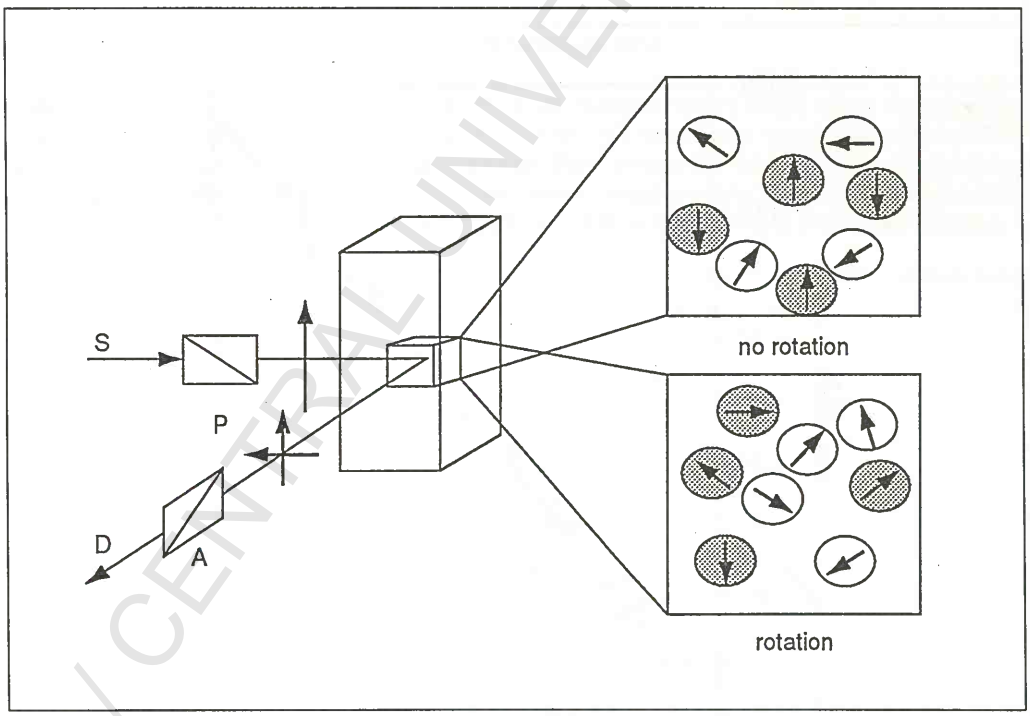


Figure 7.7.12 Depolarization of fluorescence indicates rotation of the chromophore. Monochromatic radiation from the source (S) has all but the vertically polarized electric vector removed by the polarizer (P). This is absorbed only by those molecules (see Fig. 7.7.5) in which the transition dipole of the chromophore is aligned vertically. In the case where these molecules do not rotate appreciably before they fluoresce ("no rotation"), the same molecules will fluoresce (indicated by shading) and their emitted radiation will be polarized parallel to the incident radiation. The intensity of radiation falling on the detector (D) will be zero when the analyzer (A) is oriented perpendicular to the polarizer. In the case where the molecules rotate significantly before fluorescence takes place, some of the excited chromophores will emit radiation with a horizontal polarization ("rotation") and some with a vertical polarization. Finite intensities will be measured with both parallel and perpendicular orientations of the analyzer. The fluorescence from the remainder of the excited molecules will not be detected. The heavy arrows on the left of the diagram illustrate the case where there is rotation.

FCU IAS / CENTRAL UNIVERSITY LIBRARY

(Fig. 7.7.12). The emitted radiation is now said to be depolarized. The polarization (P) can be quantitated as

$$P = (I_{par} - I_{perp})/(I_{par} + I_{perp})$$

where I_{par} and I_{perp}, respectively, are the intensities measured parallel and perpendicular to the original plane of polarization, usually vertical.

For chromophores that are part of small molecules, or that are located flexibly on large molecules, the depolarization is complete—i.e., $P = 0$. A protein of, say, $M_r = 25$ kDa however, has a rotational diffusion coefficient such that only limited rotation occurs before emission of fluorescence and only partial depolarization occurs, measured as $1 > P > 0$. The depolarization can therefore provide access to the rotational diffusion coefficient and hence the asymmetry and/or degree of expansion of the protein molecule, its state of association, and its major conformational changes. This holds provided that the chromophore is firmly bound within the protein and not able to rotate independently. Chromophores can be intrinsic tryptophan or extrinsic, covalently bound fluorophores—e.g., the dansyl (5-dimethylamino-1-naphthalenesulfonyl) group. More detailed information can be obtained from time-resolved measurements of depolarization, in which the kinetics of rotation, rather than the average degree of rotation, are measured. For further details, see Lakowicz (1983) and Campbell and Dwek (1984).

Although potentially powerful, the interpretation of fluorescence depolarization results requires extreme caution.

Table 7.7.1 provides a summary guide for interpreting results obtained using the techniques described in this unit.

In conclusion, given a protein in which tryptophan is partially or wholly buried, quantitative estimates of quenching by different quenchers provide a further fingerprint that is specific to the particular conformation of a recombinant protein and which, combined with other fluorescence, circular dichroism, and gel electrophoresis data, forms the basis for specific identification and quality assessment.

COMMENTARY

Background Information

Fluorescence depends on the atomic environment of the tryptophan and tyrosine residues and on their accessibility to solvent. Although the structural information is limited strictly to the immediate neighborhood of the relatively few aromatic residues, the cooperative nature of protein conformations means that changes in aromatic fluorescence frequently reflect more widespread changes in conformation. Similarly, measures of solvent accessibility obtained from fluorescence quenching reflect changes in the folding and packing of nonaromatic residues in a very sensitive way. Along with its sensitivity to small changes in conformation, fluorescence offers the additional advantage of requiring very small amounts of protein, in favorable cases down to tens of micrograms.

The method is less powerful in characterizing proteins lacking tryptophan, for reasons mentioned earlier (see Support Protocol), and does not apply at all to the rare proteins that lack both fluorophores. Near-UV circular dichroism measurements (*UNIT 7.6*) can still be of considerable value for proteins containing tyrosine and, in both cases, both far-UV circular dichroism and urea-gradient gel electrophoresis (*UNIT 7.4*) provide powerful contributions to characterization.

Extrinsic fluorescent probes can provide additional useful information about the integrity of protein conformation. 8-{2-[(iodoacetyl)ethyl]amino} naphthalene sulfonic acid (ANS) fluoresces only weakly in free solution but strongly when bound to a nonpolar surface. Originally used as a competitive probe to monitor fatty-acid binding to albumin, it has been widely used in studies of protein conformation—in partially denatured states at equilibrium as well as in transient intermediates that accumulate during protein folding (reviewed

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

by Christensen and Pain, 1994). Fully folded native proteins in general bind only limited amounts of ANS, although these are sufficient to characterize the surface hydrophobicity of a number of proteins (Cardamone and Puri, 1992). Fully unfolded proteins do not bind ANS. Partially folded or partially unfolded proteins will, however, bind ANS—indicating the presence of associated nonpolar groups that constitute a surface site, or the existence of relatively loosely packed hydrophobic cores into which ANS can intercalate. This and sim-

Table 7.7.1 Summary of the Interpretation of Fluorescence Spectra of Proteins[a,b]

1. All fluorescence of a protein is due to tryptophan, tyrosine, and phenylalanine, unless the protein contains another fluorescent component.
2. Maximum fluorescence of tryptophan occurs at lower wavelengths (<350 nm) and has greater intensity the lower the polarity of its environment.
   a. If λ_{max} is located at lower wavelengths when the protein is in a polar solvent—e.g., water—the tryptophan must be buried in a nonpolar environment.
   b. If λ_{max} is shifted to shorter wavelengths when the protein is transferred to a nonpolar medium, either the tryptophan is on the surface of the protein or the solvent induces a conformational change that causes it to be exposed to the solvent.
3. If a substance known to be a quencher quenches tyrosine and/or tryptophan fluorescence in a protein, these residues must be on the surface. If it fails to do so, there are several possible reasons:
   a. The residue may be buried in the interior of the protein.
   b. The residue may be in a crevice whose dimensions are too small for the quencher to enter.
   c. The residue may be in highly charged surroundings whose charge may repel a quencher with the same charge. In such a case, a quencher of the opposite charge or no charge would be expected to quench the fluorescence.
4. If a substance that does not affect the quantum yield of the free amino acid affects the fluorescence of a protein, it probably does so by inducing a conformational change in the protein.
5. If tryptophan or tyrosine are in a polar environment, their quantum yield decreases with increasing temperature, whereas in a nonpolar environment there is little change. Deviation from a monotonic decrease with increasing temperature signifies a change in polarity of the tryptophan environment, the sign of the change indicating the direction of change in polarity. Either indicates a conformational change in the protein.
6. Tryptophan fluorescence is quenched by protonated acidic groups, either in the neighboring environment or covalently attached as an α-carboxyl group. If the fluorescence changes with a particular pK, then a residue with that pK must be close to the tryptophan. Such an interpretation only holds if evidence is provided that no pH-induced conformational change taking place.
7. If the binding of a substance is accompanied by quenching of tryptophan fluorescence, either there is a resulting conformational change or tryptophan is in or very near the binding site. Furthermore, because a decrease in polarity causes a shift to shorter wavelengths, a decrease in λ_{max} on binding indicates that water is excluded in the complex.
8. If the absorbance spectrum of a small molecule overlaps the emission spectrum of tryptophan and if the distance between them is small, quenching will be observed. If binding of such a molecule to a protein quenches tryptophan fluorescence, tryptophan must be in or near (<2 nm away from) the binding site.

[a]Adapted from Freifelder (1982). Reproduced with permission of W.H. Freeman, New York.

[b]In making these interpretations, two important facts must be borne in mind. First, fluorescence is so sensitive to environmental factors that alternative interpretations must always be considered. Second, if a protein contains more than one tryptophan, each may have a different quantum yield. For this reason the absolute magnitude of changes cannot be used to determine the fraction in a given environment.

CU IAS/CENTRAL UNIVERSITY LIBRARY

ilar probes are mainly used as empirical indicators of conformational change, as the detailed mechanisms of interaction are not well defined. Fluorescence energy transfer, which can be observed when the absorbance spectrum of a fluorescent probe such as the dansyl group overlaps the emission spectrum of tryptophan, provides a means of estimating distances between groups, and hence of conformation and conformational change. An example where this approach is used alongside ANS fluorescence to follow protein folding is reported by Agashe et al. (1995).

In the present context, fluorescence is an important tool for the rapid comparison of a sample of recombinant wild-type or mutant protein with a standard sample of the recombinant or the authentic protein, as part of the process of characterizing and authenticating a recombinant protein. Coupled with measurements of the depolarization of fluorescence, the method can provide information on the dynamics of a protein and on the change in dynamics in response to, for example, binding of ligands.

Critical Parameters

Recording a spectrum is basically very simple, but the critical element is to avoid the production of artifacts. For this reason, it has to be emphasized that scrupulous attention to the cleanliness of cells and of the cell compartment of the spectrofluorimeter is essential, as is attention to the clarification of solutions and buffers (see Strategic Planning).

It is important also to make intelligent choices in the selection of instrumental parameters to optimize resolution and signal-to-noise ratio (see Basic Protocol 1).

Troubleshooting

One of the main causes of artifacts in the spectroscopy of proteins is the frequent tendency of proteins to aggregate, either during the folding step in recovery from inclusion bodies or as a result of the delicate balance of solubility of many proteins. For this reason, it is important to pay strict attention to monitoring aggregation (see discussion of Clarification of Solutions in Strategic Planning).

Literature Cited

Agashe, V.R., Shashtry, M.C.R., and Udgaonkar, J.B. 1995. Initial hydrophobic collapse in the folding of barstar. *Nature* 377:754-757.

Cardamone, M. and Puri, K. 1992. Spectrofluorimetric assessment of the surface hydrophobicity of proteins. *Biochem. J.* 282:589-593.

Campbell, I.D. and Dwek, R.A. 1984. Biological Spectroscopy. Benjamin/Cummings, Menlo Park, Calif.

Christensen, H. and Pain, R.H. 1994. The contribution of the molten globule model. *In* Mechanisms of Protein Folding (R.H. Pain, ed.) pp. 55-79. Oxford University Press, Oxford.

Eftink, M.R. and Ghiron, C. 1981. Fluorescence quenching studies with proteins. *Anal. Biochem.* 114: 199-227.

Hlodan, R. and Pain, R.H. 1994. Tumour necrosis factor is in equilibrium with a trimeric molten globule at low pH. *FEBS Lett.* 343:256-260.

Lakowicz, J.R. 1983. Principles of Fluorescence Spectroscopy. Plenum Press, New York.

Schmid, F.X. 1989. Spectral methods of characterizing protein conformation and conformational changes. *In* Protein Structure (T.E. Creighton, ed.) pp. 251-285. IRL Press, Oxford.

Varley, P.G., Dryden, D.T., and Pain, R.H. 1991. Resolution of the fluorescence of the buried tryptophan in yeast 3-phosphoglycerate kinase using succinimide. *Biochim. Biophys. Acta* 1077:19-24.

Key Reference

Lakowicz, J.R. 1983. See above.

An essential text, providing a thorough, but clear and readable description of the practice and interpretation of fluorescence spectroscopy, with particular reference to proteins.

Contributed by Roger H. Pain
Jožef Stefan Institute
Ljubljana, Slovenia

BCU IASI CENTRAL UNIVERSITY LIBRARY

Light Scattering

Light scattering methods can provide information about the native molecular weight, oligomeric composition, and gross conformation of a protein in solution. These methods are particularly well suited for studying large oligomeric systems or glycoproteins and can be used to characterize much larger structures involving protein such as viruses and even bacterial spores (Harding, 1997). Light scattering techniques are not so well suited for characterization of smaller protein systems (where the molecular weight, M, is <30,000 Da); in these cases other methods such as analytical ultracentrifugation (*UNIT 7.5*) and solution X-ray scattering are more suitable. All light scattering measurements on solutions of proteins and protein assemblies are based on the principle of analyzing the intensity of light scattered by the solution (Fig. 7.8.1), either in terms of the time-averaged intensity ("classical" or "static" light scattering) or intensity fluctuations with time ("dynamic" or "quasielastic" light scattering) at a given angle or series of angles.

There are three types of "static" light scattering experiment:

1. Turbidimetry, which is simple but gives only crude molecular-weight estimates for large assemblies;
2. Low-angle light scattering, which is also simple and gives molecular-weight and molecular-weight-distribution information;
3. Multiangle light scattering, which gives more reliable molecular-weight and molecular-weight-distribution information and, for proteins of molecular weight at least ~50,000 Da, solution conformation information.

There are also three types of "dynamic" light scattering measurements:

1. Fixed-angle (90°-angle) measurements, which are simple and give an estimate for the translational diffusion coefficient and an idea of sample polydispersity for approximately globular macromolecules;
2. Variable-angle measurements, which are less simple, give more reliable estimates for translational diffusion coefficient and sample polydispersity, and can in some circumstances yield an estimate for rotational diffusion coefficients;

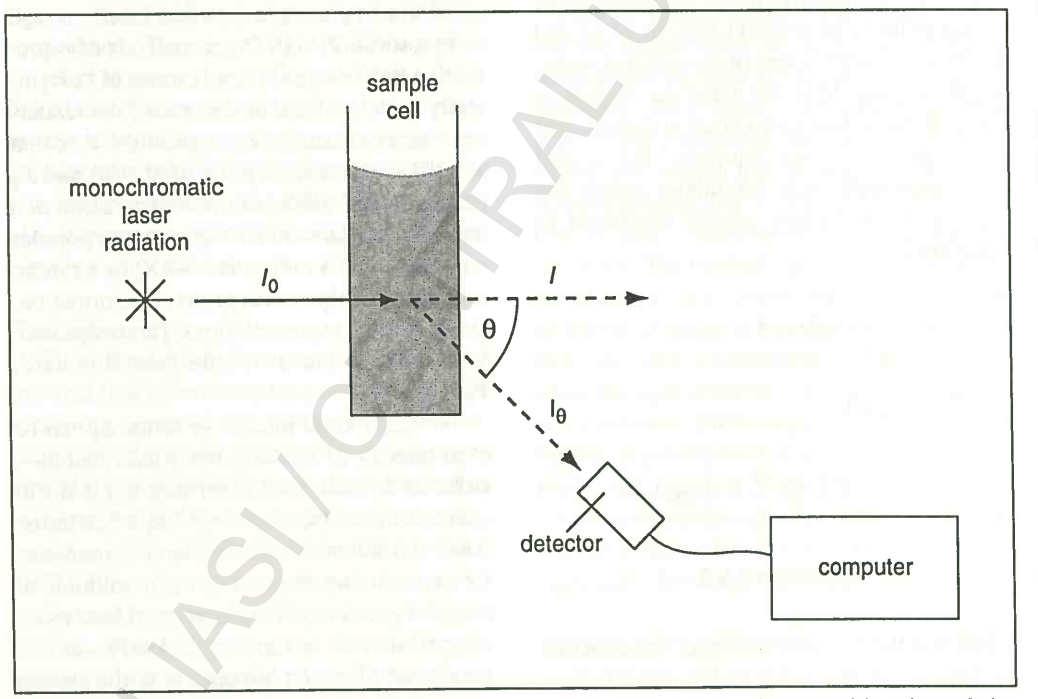


Figure 7.8.1 Light scattering studies on biomolecular solutions involve consideration of the relationship between the time-averaged scattered light intensity, $I(\theta)$, with the incident intensity, I_0, and angle of detection, θ ("static light scattering") or of the rapid fluctuations of the scattered intensity, $I(\theta)$, with time, τ ("dynamic light scattering"). Turbidimetry involves static light scattering measurements on the relation between I_0 and transmitted light intensity, I (at zero angle) alone.

Contributed by Stephen E. Harding and Kornelia Jumel
Current Protocols in Protein Science (1998) 7.8.1-7.8.14
Copyright © 1998 by John Wiley & Sons, Inc.

BU IASI / CENTRAL UNIVERSITY LIBRARY

3. Electrophoretic light scattering measurements ("ELS"), popularly used for studying colloid solubility.

STATIC LIGHT SCATTERING ANALYSIS OF PROTEIN SOLUTIONS

Basic Theory

The basic equation for the angular dependence of light scattered from a solution of proteins or protein assemblies is the Debye-Zimm relation (Zimm, 1948), shown in Equation 7.8.1,

$$\frac{Kc}{R_\theta} = \frac{1}{MP(\theta)}(1 + 2A_2c + \ldots)$$

Equation 7.8.1

In this equation, A_2 (ml mol/g^2) is the thermodynamic (2nd virial) nonideality coefficient. R_θ is the Rayleigh excess ratio (the ratio of the intensity, I_θ, of excess light scattered compared to pure solvent) at an angle θ to that of the incident light intensity, I_0 (a correction term is necessary if unpolarized light is used but not necessary if lasers are used). K is an experimental constant dependent on the square of the solvent refractive index, the square of the refractive index increment (*dn/dc* in ml/g), and the inverse fourth power of the incident wavelength, λ (cm). M is the molecular weight in Da, c is the solute concentration (g/ml), and $P(\theta)$ is the form factor. Equation 7.8.1 is valid if the proteins/protein assemblies satisfy the Rayleigh-Gans-Debye criteria illustrated in Equation 7.8.2:

$$\left|\frac{n}{n_0} - 1\right| << 1$$

and

$$\left(\frac{4\pi n_0}{\lambda_0}\right) d \left|\frac{n}{n_0} - 1\right| << 1$$

Equation 7.8.2

where n is the refractive index of the solution, n_0 is the refractive index of the solvent, λ_0 is the incident wavelength (in vacuo), and d is the maximum dimension of the particle. The form factor $P(\theta)$ can also be given to a good approximation by Equation 7.8.3,

$$\frac{1}{P(\theta)} \approx \left[\frac{1 + 16\pi R_\theta^2}{3\lambda^2} \sin^2\left(\frac{\theta}{2}\right)\right]\left(\frac{1}{M} + 2A_2c\right)$$

Equation 7.8.3

R_g is extensively referred to as the "radius of gyration" of the macromolecule and c is the solute concentration (g/ml). If the solute is heterogeneous, M (g/mol = Da) will be a weight average, M_W, and R_g a z-average. Equation 7.8.1 is generally a good representation for particles whose maximum dimension is between $\lambda/20$ and λ. Since λ is typically ~600 nm, this covers all proteins and protein assemblies up to the sizes of filamentous viruses. For larger particles, much more complex representations are necessary.

For particles of dimension $<\lambda/20$ ($M <$ 50,000 Da) the angular term in Equation 7.8.1 is small. No angular-dependence measurements are necessary to obtain M (although R_g cannot be obtained; if this is needed then X-ray or neutron solution scattering measurements need to be employed instead). For particles of dimension $>\lambda/20$, a double extrapolation to zero angle and zero concentration is necessary. This is usually performed on a grid-like plot referred to as a "Zimm plot" or via measurement at a single angle assumed small enough so that $\sin^2(\theta/2) \sim 0$. Other methods of representing the data have been in terms of "disymmetry": $z(\theta)$ (defined as the ratio of the scattering intensity at an angle θ, typically 45°, to that at 180° − θ) versus θ plots. Both $z(\theta)$ and R_g provide useful guides to the conformation of a macromolecule, with $z(\theta)$ being more popular with linear DNA molecules and R_g being more suited for descriptions of protein conformation. Other useful representations for nonprotein systems are also available (see Burchard, 1992).

For fairly rigid protein systems, R_g can be used directly to model gross conformation—either as an additional parameter to the diffusion coefficient (see Dynamic Light Scattering Analysis) and other hydrodynamic parameters for representing the structure in solution of complex protein systems in terms of bead modeling (Garcia de la Torre et al., 1997)—or as a parameter, after combination with the second virial coefficient A_2 and a parameter from solution viscometry known as the intrinsic viscosity, for representing the triaxial structure of a protein (Harding et al., 1997).

The principal useful parameters to be derived from static light scattering measurements

BCU IASI / CENTRAL UNIVERSITY LIBRARY

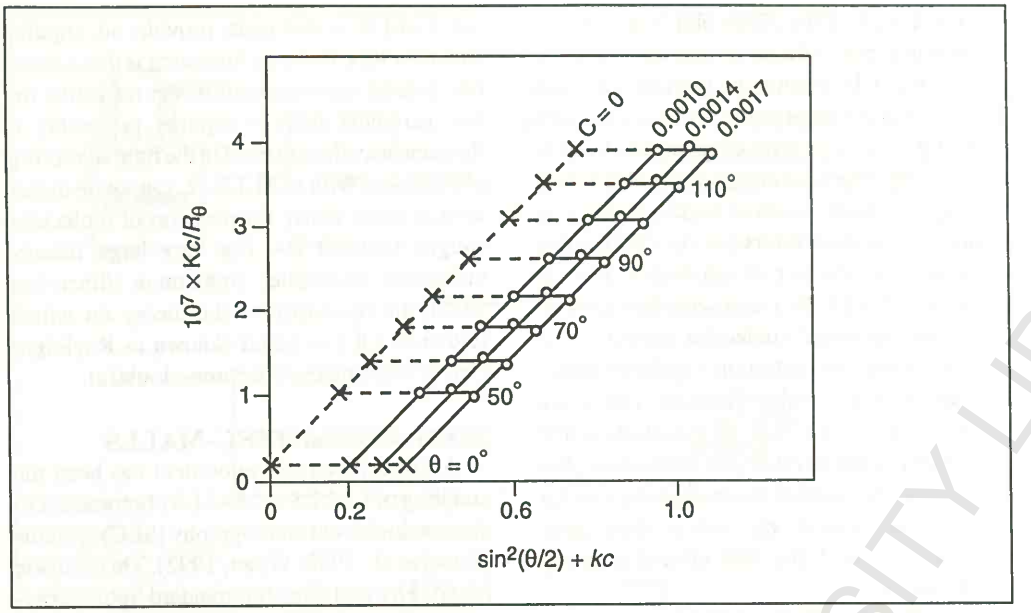


Figure 7.8.2 Zimm biaxial plot for a (diptheria) toxin-antibody aggregate. The arbitrary scaling constant, *k*, equals 200 ml/g. The molecular weight (M_w) as calculated from the (reciprocal) common intercept of the $c = 0$ and $\theta =$ lines is equal to ~78×10^6 Da. Data from Johnson and Ottewill (1954) and Johnson (1993).

are thus M, R_g, and to a lesser extent A_2. M can be obtained to a reasonable accuracy (usually to within 5%). The extraction of R_g is more difficult, requiring considerable care in the form of the angular extrapolation, and becomes even more difficult as the lower limit of $\lambda/20$ is approached. The use of light scattering photometers incorporating a flow cell that can be linked directly on-line to size-exclusion chromatography columns is becoming increasingly popular, particularly for the characterization of polydisperse systems (the hallmark, for example, of many glycoproteins), and as an on-line sample clarification system.

Turbidimetry

Turbidimetry involves the measurement of the total loss of intensity by a solution through scattering, summed over the entire angular intensity envelope and compared with the intensity of the incident radiation. It can be used to measure the molecular weights of protein assemblies of $M > 10^5$ Da (Bahls and Bloomfield, 1977). It is a type of measurement that can be performed on a good-quality spectrophotometer (i.e., one whose detector does not accept appreciable amounts of scattered light). Measurements have to be made at wavelengths away from the influence of absorption maxima.

Low-Angle Light Scattering (LALLS)

As θ approaches 0, Equation 7.8.1 becomes Equation 7.8.4:

$$\frac{Kc}{R_\theta} = \frac{1}{M} + 2A_2c$$

Equation 7.8.4

Low-angle light scattering photometers permit the measurement of Kc/R_θ at one fixed, small scattering angle θ (usually <8°) at which Equation 7.8.4 is taken to be valid (see Jumel et al., 1992). Although the angle used is assumed low enough that no angular correction of the scattering data is required, an extrapolation to zero concentration of Kc/R_θ may be necessary. The method can thus provide values for M and A_2 of a system, but not for R_g since no record is made of the angular dependence of Kc/R_θ. At low concentration, and especially when the light scattering detector is on-line to size-exclusion chromatography columns, the further approximation that $2A_2c \sim 0$ can be made, and hence M is simply $\sim R_\theta/Kc$. Since lasers are now routinely used as the light source, this technique takes the popular acronym of "LALLS" (low-angle laser light scattering).

Multiangle Light Scattering (MALLS)

Multiangle light scattering measurements are based on the full Debye-Zimm equation (Equation 7.8.1). Performing measurements at multiple angles permits extrapolation of the ratio Kc/R_θ to zero $\sin^2(\theta/2)$, which, together with an extrapolation to zero concentration,

POLISH ACADEMY OF SCIENCES CENTRAL UNIVERSITY LIBRARY

forms the basis of the Zimm plot (Fig. 7.8.2), a method that can yield M, A_2 and R_g.

Values for A_2 returned are typically between 10^{-5} ml mol/g^2 (for globular proteins) and 10^{-3} ml mol/g^2 (for large glycoconjugates). A warning is in order here—a common misconception is to ignore these values as small numbers. In actuality, it is the product $2 \times A_2 \times M \times c$ that manifests the influence of nonideality. The expression $1/(1 + 2A_2Mc)$ represents the factor by which an "apparent" molecular weight measured at a finite concentration c underestimates the true or "ideal" value (Tanford, 1961). An opposing effect to that of thermodynamic nonideality is that of reversible interaction phenomena (as represented for example by a molar dissociation constant, K_d), and in some cases ("pseudoideality") the two effects approximately cancel.

As with LALLS, lasers are now routinely used as the light source, and hence the acronym "MALLS" (multiangle laser light scattering) has now been adopted. Besides the potential for extracting R_g, a more important advantage of MALLS over LALLS is that the angular extrapolation permits identification and avoidance of any spurious results at the lowest angles. Plots of Kc/R_θ are only linear over a limited range of angles. For globular proteins of $M <$ 50,000 Da (corresponding to a maximum dimension ~$\lambda/20$), the angular dependence of K_c/R_θ will be negligible, permitting a more accurate determination of M but of course excluding the possibility of measuring R_g. As with LALLS, at low concentration, the further approximation that $2A_2c \sim 0$ can be made; hence Equation 7.8.1 reduces to

$$\frac{Kc}{R_\theta} = \frac{1}{MP(\theta)}$$

Equation 7.8.5

and where there is no angular dependence, this reduces further to $M \sim R_\theta/Kc$, as noted above.

Limitations of Static Light Scattering Methods

The principal limitation for all these "static" light scattering measurements is the need for sample clarification to avoid dust and supramolecular aggregates (see Clarification of Solutions and Scattering Cells). This is especially serious if solutions of proteins of M <50,000 Da are being studied, and also if LALLS is used, since large-particle contamination effects become disproportionally larger at low angles and the LALLS instruments provide no angular check for this. Another limitation is that a separate precise measurement of the refractive index increment dn/dc is required, preferably at the same wavelength used in the light scattering photometer. With MALLS, R_g cannot be measured as noted above for particles of molecular weight <50,000 Da. For very large macromolecular assemblies (maximum dimension >$\lambda/2$), the reliability of the theory on which Equation 7.8.1 is based (known as Rayleigh-Gans-Debye theory) becomes doubtful.

SEC-LALLS and SEC-MALLS

A revolutionary development has been the coupling of LALLS or MALLS photometers to size-exclusion chromatography (SEC) systems (Jumel et al., 1992; Wyatt, 1992). This is made possible by replacing the standard light scattering cuvette with a flow cell that can be coupled downstream from an HPLC pump and size-exclusion chromatography column(s), and upstream from a (UV-absorption or refractive index) concentration detector (Fig. 7.8.3). This has the double advantage of providing an online clarification system to remove supramolecular contamination and allowing fractionation of polydisperse materials prior to light scattering analysis. However, care must still be taken, particularly against the spurious shedding of column material. The coupled SEC systems are referred to as "SEC-LALLS" or "SEC-MALLS."

Samples for Analysis in Static Light Scattering

Solutions should be dialyzed against an appropriate buffer of defined pH and ionic strength, I. Normally an ionic strength of at least 50 mM is needed to suppress the contribution of molecular charge on the protein to the nonideality coefficient (A_2 in Equation 7.8.1) although, of course, the I chosen will obviously depend on its effect on the protein.

The concentration of protein required depends on: (1) its molecular weight, since the scattering signal is approximately proportional to the product of concentration and molecular weight; (2) the output of the laser; and (3) the clarity of the solutions, since substantial material may be lost on filtering or from the column in SEC-LALLS or SEC-MALLS. For scrupulously clean solutions, a 5-mW laser for a loading concentration of ~3-mg/ml is sufficient for proteins of molecular weight ~40,000 Da. For smaller proteins, a proportionally higher con-

BCU IAS CENTRAL UNIVERSITY LIBRARY

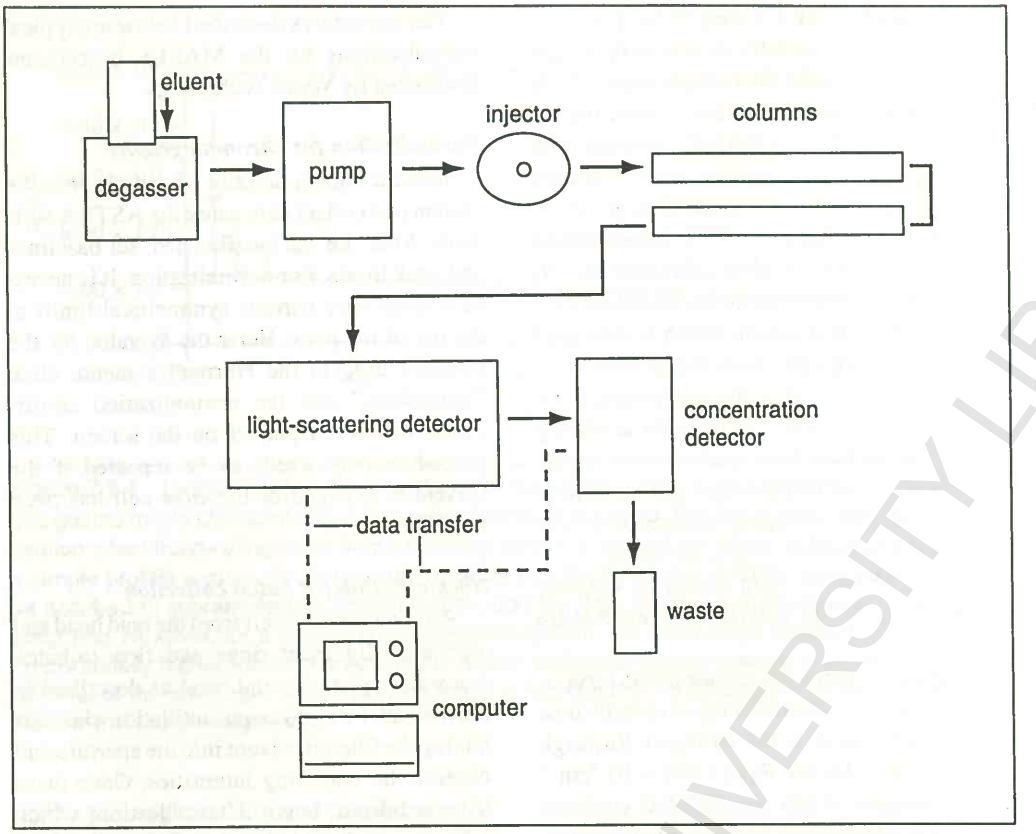


Figure 7.8.3 Experimental setup for SEC-MALLS (see Wyatt, 1992).

centration and/or higher laser power is required.

With regard to volume requirements, if a flow cell arrangement is used (linked, e.g., to an SEC system), loading volumes can be as low as 0.1 ml. For use with scintillation vials (which are nowadays used instead of fluorimeter-type cuvettes), ≥3 ml of solution is required. Square cuvettes, of course, prevent measurements in the angular part of the scattering envelope near the corners of the cuvette; if cylindrical cuvettes are employed, then small diameters (<2 cm) are to be avoided because of extraneous scattering or reflections from the glass walls, although large-diameter cuvettes can be expensive in terms of quantity of solution required. Flow cells are now preferred, and a typical experimental set up would have MALLS coupled to an SEC, but with a separate injection port (via a filter or guard column) if the SEC separation is not needed.

Clarification of Solutions and Scattering Cells

For molecular weights <200,000 Da, the most serious experimental problem has been that of clarification. All traces of dust and supramolecular aggregates have to be removed since experiments on incompletely purified material are "not useful" (Johnson, 1993), and clarification can lead to a significant loss of material. Contaminating particles can be removed by ultracentrifugation and ultrafiltration.

If scintillation vials are used, these too must be scrupulously clean. They should be cleaned and rinsed in double-distilled water, then dried and covered with aluminum foil to prevent any dust from entering the vials. It is also extremely important that there are no scratches or fingerprints or other marks on the outside of the vial. Sample preparation for batch work should, if possible, be carried out in a flow cabinet to prevent any dust from entering the solutions.

If a flow-cell device is used for clarification (Sanders and Cannell, 1980), this too will frequently need to be removed and cleaned using detergent, acetic acid, and a final necessary rinse with ultrafiltered water. An acetone reflux may also be occasionally necessary. When attached to an SEC column, prior to solution injection and after all the solution has been eluted, the solvent eluant must be monitored for any shedding from the column material. When the solution is injected, this needs to be done via a Millipore filter of appropriate size.

BOSTON UNIVERSITY LIBRARY

Calibration (LALLS and MALLS)

The quantities directly measured by a light scattering photometer are voltages and not light scattering intensities. For this reason, the instrument has to be "calibrated"—usually with a strong Rayleigh scatterer with a known Rayleigh ratio such as toluene (Stacey, 1956; Johnson and McKenzie, 1977). Measurement of the scattered and incident light intensities (I_θ and I_0, respectively) then facilitates calculation of the calibration constant, which is then used to calculate Rayleigh ratios for sample solutions from the output of the instrument. Calibration does *not* mean that the light scattering measurements have to be made relative to protein standards as in calibrated gel chromatographic methods. A typical calibration procedure would include the following steps.

1. Switch on the light scattering detector and laser at least 1 hr before measurements are made.
2. Make sure that the sample cell is scrupulously clean and the calibrating solvent (in most cases toluene as it is the strongest Rayleigh scatterer with a known R_θ of 1.406×10^{-5}cm^{-1} at a wavelength of 633 nm) is HPLC grade.
3. Inject the solvent into the cell via a membrane filter (preferably ≤0.2 μm pore size).
4. Measure the scattered light intensity and calculate the instrument calibration constant by means of the dedicated software (e.g., ASTRA, Wyatt Technology), using the known Rayleigh ratio of the solvent.

Normalization (MALLS)

For simultaneous multiangle detection in MALLS photometers of the type now commonly available, the detectors have to be "normalized" to allow for the different scattering volumes as a function of angle and the differing responses of the detectors. This is normally achieved using a solution of molecules that scatter light isotropically (i.e., with equal intensity in all directions)—as is the case when the diameter of the sample molecule is less than ~λ/20 of the incident wavelength. The best molecules of this kind for use in a chromatography system are either a polystyrene standard in toluene or tetrahydrofuran (THF) with molecular weight of ~30,000 Da or a pullulan or dextran standard of 20,000 to 30,000 Da in aqueous solution; the R_g of these molecules is ~5 nm. In the batch mode—i.e., with the sample in a vial–10 mg/ml solutions of 4000-Da polystyrene in toluene or THF, or 5000 Da pullulan in water work well; these standards have R_g's of ~2 nm.

The procedures described below are typical normalizations for the MALLS instruments distributed by Wyatt Technology.

Normalization for chromatography

Inject a known amount of sample into the system and collect data using the ASTRA software. After the run has finished, set baselines and peak limits. For normalization, it is necessary to set very narrow, symmetrical limits at the top of the peak. Enter the R_g value for the standard used in the Normalize menu, click "normalize," and the normalization coefficients will be displayed on the screen. This procedure only needs to be repeated if the solvent is changed or the flow cell has been cleaned.

Normalization for batch collection

Remove the flow cell from the read head and assemble alignment rings and flow-to-batch conversion plate in read head as described in the manual. First, place a scintillation vial containing the filtered solvent into the aperture and observe the scattering intensities. Once these have stabilized, begin data collection, which will stop after a preset volume. Save the data. Next, insert the vial containing the sample for normalization measurement and again wait for signals to stabilize. Append the file containing the values for the solvent. After collection has stopped, measure the scattering for the solvent again. Set a baseline and peak limits that cover most of the plateau corresponding to the normalization standard. Go to the Normalization menu and enter the concentration and R_g of the normalization standard. Click "normalize" and the normalization coefficients will appear on the screen.

Evaluation of Molecular Weight (LALLS)

After measurement of R_θ, the molecular weight, M, can be calculated from this value, the concentration, the experimental constant K, and a subsequent extrapolation to $c = 0$ (Equation 7.8.4) if the nonideality term A_2c is significant. The concentration, c (from, e.g., UV absorbance or refractometry) should be known as accurately as possible.

Evaluation of Molecular Weight and Radius of Gyration (MALLS).

If the nonideality term A_2c (and of course any higher-order terms) is insignificant (i.e., if the measurements can be performed at low enough concentration), then only an extrapola-

BCU IASTER CENTRAL UNIVERSITY LIBRARY

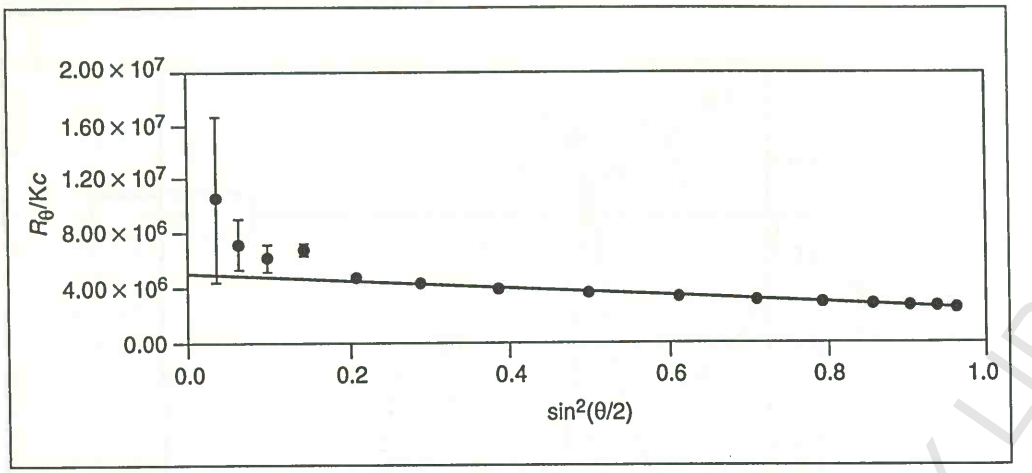


Figure 7.8.4 Debye plot (angles from 21.7° to 58.3°) for a heavily glycosylated protein system (pig gastric mucin glycoprotein). A first-order (linear) fit is shown. This figure illustrates the need for caution when interpreting static light-scattering data: the upward curvature at the lower angles for a simple protein system (or downward curvature in a Kc/R_θ versus $\sin^2(\theta/2)$ plot) would normally be ascribed to supramolecular contamination, and the data points would be ignored or given low weighting. However, for a genuinely polydisperse system, the lower-angle measurements contain proportionally higher information about the higher-molecular-weight species in a distribution, and removal of the low-angle data and/or choice of a first-order fit could bias the distribution towards the low-molecular-weight end.

tion to zero angle and not zero concentration is necessary (see Equation 7.8.5). Sometimes the inverse form of Equation 7.8.4 is used, known as a "Debye plot" of R_θ/Kc versus $\sin^2(\theta/2)$ (Wyatt, 1992). Although this invokes a further approximation, the ordinate axis does correspond to an apparent molecular-weight axis (Fig. 7.8.4). It must be noted that Figure 7.8.4 is a rather extreme example for a highly glycosylated and polydisperse system. Figure 7.8.4 also shows that care must be taken in choosing the appropriate fit, and that there is a proportionally larger effect at low angles of supramolecular contaminants. The order of the fit chosen (usually linear or quadratic) has a more dramatic effect on the (limiting) slope of such plots and hence on the value for R_g returned. This feature needs to be borne in mind when conclusions on macromolecular conformation based on the relationship between M and R_g are being drawn, particularly using the SEC-MALLS devices considered below.

If the nonideality term is significant, and an assumed value for A_2 cannot be taken as zero (which can in fact be predicted from the triaxial shape of the protein; Harding et al., 1997)—and charge effects are not significant, then an additional extrapolation to zero concentration is necessary. Both angular and concentration extrapolations are usually done on the same plot, known as a Zimm plot (Fig. 7.8.2), which can yield M from the reciprocal of the common intercept, and estimates for A_2 (from the $\theta = 0$, c extrapolation line) or R_g [from the $c = 0$, $\sin^2(\theta/2)$ extrapolation line]. However it cannot be overstressed that the R_g returned can be very sensitive to the order of extrapolation used. An arbitrary (positive or negative) constant k is used on the abscissa to scale the data points better.

Evaluation of Molecular Weight Distribution (SEC-LALLS) and (SEC-MALLS)

Absolute molecular weight distributions (Fig. 7.8.5) of mixed protein systems or glycosylated protein systems such as mucins, glycoproteins, or glycosaminoglycans may be obtained with either of the above techniques.The sample is injected via the injection valve and separated by the column system. Effluent from the column(s) is monitored by light scattering and concentration detectors. Once the peak area from the chromatogram has been chosen, the concentration and R_θ at each data point within this area are known from light scattering and concentration detectors, respectively, and the molecular weight (M) at each of these points is calculated using Equation 7.8.4 (for SEC/LALLS) or Equation 7.8.5 (for SEC/MALLS). Molecular weight averages for the whole of the peak area may then be calculated. In addition, a so-called "calibration plot" of molecular weight versus elution vol-

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

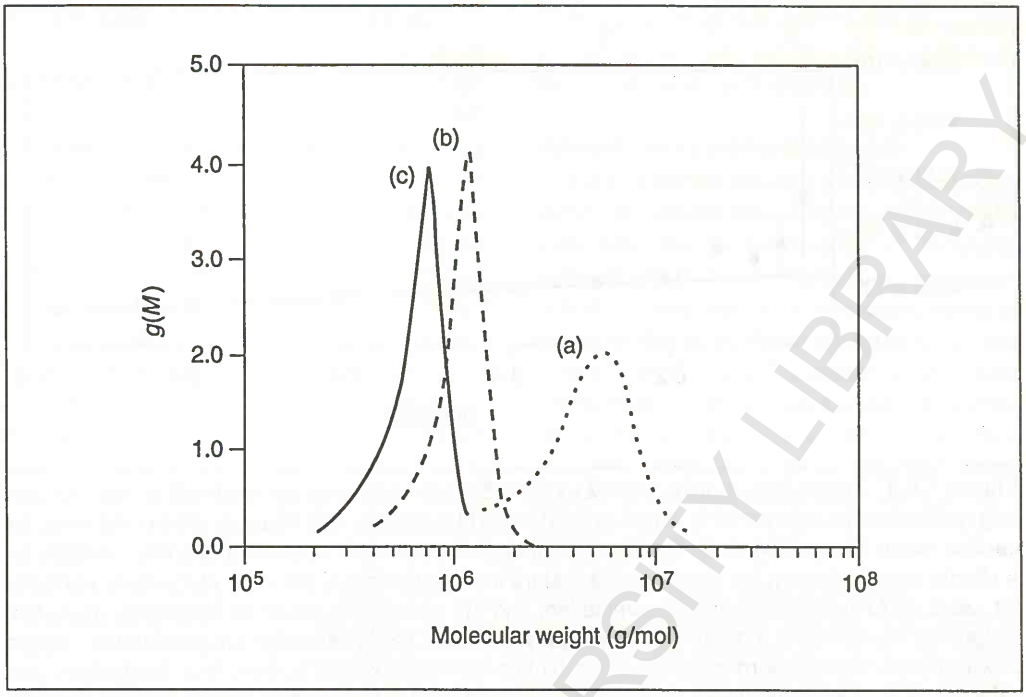


Figure 7.8.5 SEC-MALLS molecular-weight distribution for colonic mucin glycoprotein (peak a) and its thiol-reduced (peak b) and papain digested (peak c) forms (from Jumel et al., 1997). The quantity g(M) is in arbitrary units.

ume can be constructed, thus enabling the elution volume versus concentration plot to be converted to a molecular weight distribution. Such molecular weight distributions thus found will be absolute in the sense that, unlike with conventional SEC, calibration standards of known *M* are not required. This is particularly useful not only for protein mixtures but for heavily glycosylated systems because of their polydispersity and the difficulty in obtaining standards of the appropriate (often uncertain) conformation.

For a more detailed account of the theory the reader is referred to Yau et al. (1979).

DYNAMIC LIGHT SCATTERING ANALYSIS OF PROTEIN SOLUTIONS

Basic Theory

Whereas static light scattering is concerned with the time-averaged scattering intensity properties, dynamic light scattering is concerned with time fluctuations in intensity caused by motions of macromolecules and macromolecular assemblies. Whereas lasers are highly desirable for static light scattering because of their high intensity, collimation, and monochromaticity—with dynamic light scattering they are mandatory because of the requirement for spatial and time ("temporal") coherence—light has to be emitted from the source as a continuous wave rather than as short bursts.

The physics underlying dynamic light scattering is rather complicated (Brown, 1993), although the basic principle and experimental setup is relatively simple (Fig. 7.8.1 and Fig. 7.8.6). Laser light is directed onto a thermostatted protein solution, and the intensity is recorded at either a single angle or multiple angles using a photomultiplier/photodetector. The moving biomolecules will "Doppler broaden" the otherwise monochromatic incident radiation. The scattered intensity recorded as the number of photons received by a detector (photomultiplier) will fluctuate because of "beating interference" of scattered waves of different but similar wavelength. The situation is analogous to the fluctuations in intensity of a radio channel caused by interference from another radio channel of very close wavelength. This "broadening" of the otherwise monochromatic incident radiation is why the technique is often referred to as "quasielastic light scattering" or QLS. The detector sends the intensity signal to a special computer called an autocorrelator, which compares or correlates the intensity at different times (hence the other name often used—photon correlation spectroscopy or PCS).

BCU IAS / PERTANIAN UNIVERSITY LIBRARY

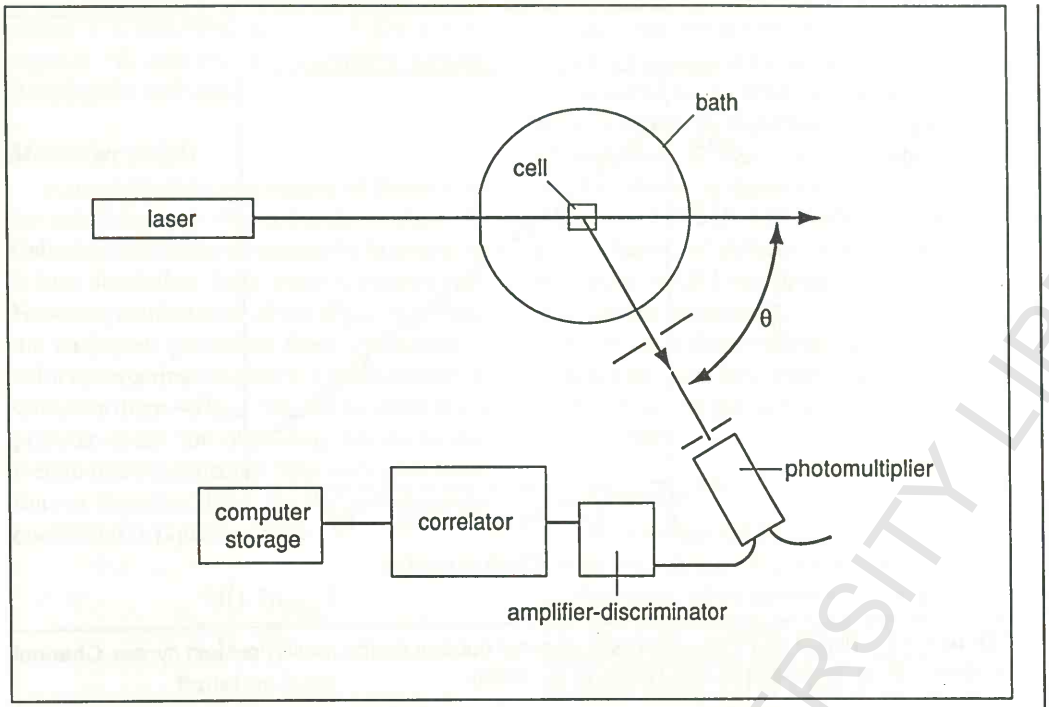


Figure 7.8.6 Schematic dynamic light scattering setup for multiangle measurement (see Johnson, 1993).

How rapidly the intensity fluctuates over short time periods or "delay times," τ (in nsec to msec depending on how mobile the scattering biomolecules/assemblies are), is represented by how a parameter known as the normalized intensity autocorrelation function—$g^{(2)}(\tau)$—decays as a function of τ. The superscript "(2)" is used to indicate that it is an intensity as opposed to an electric field or "(1)" autocorrelation function, and many data sets of $g^{(2)}(\tau)$ as a function of τ are accumulated and averaged. The amount of averaging necessary depends on the incident laser intensity and the size and concentration of the scattering biomolecule (at a given concentration, larger molecules scatter more). For globular proteins, sufficient data can usually be acquired over the time scale of one to several minutes. The calculation of the correlation function from the intensity fluctuations is performed after the signal has passed through an amplifier-discriminator via the autocorrelator. Calculation of the diffusion coefficient from the decay of $g^{(2)}(\tau)$ with τ is performed on a computer. Some of the most modern instrumentation (particularly fixed-angle photometers) have all of the units depicted in Figure 7.8.6 built into one instrument.

Analysis of how the normalized intensity autocorrelation function $g^{(2)}(\tau)$ decays as a function of τ can be used to evaluate the translational diffusion coefficient, D. For dilute systems of spherical or near-spherical (i.e., globular) biomolecules and assemblies, the variation of $g^{(2)}(\tau)$ with τ can be represented by the simple logarithmic equation

$$\ln\left[g^{(2)}(\tau-1)\right]=-2Dg^{(2)}\tau$$

Equation 7.8.6

where q is known as the Bragg wave vector whose magnitude is defined by $q = \{4\pi n/\lambda\}\sin(\theta/2)$—$n$ being the refractive index of the medium, θ being the scattering angle, and λ being the wavelength of the incident light. Thus D can be found from a plot of $\ln[g^{(2)}(\tau) - 1]$ versus τ, and Figure 7.8.7 shows an example for the motility protein dynein.

D can then be converted to standard conditions (i.e., the viscosity and temperature of water at 20.0°C) to give $D_{20,w}$ as noted above, and then extrapolated to zero concentration to give $D^0_{20,w}$. An additional extrapolation is necessary if the biomolecule is not globular: this is because at finite angles θ there will be an extra term on the right-hand side of Equation 7.8.6 deriving from rotational diffusional phenomena. This term approaches 0 as θ approaches 0; therefore true D can be measured according to Equation 7.8.6 so long as measurements of the apparent D are made at a

B. C. U. „M. Antonescu" IAŞI
CHILIE

BOURAT CENTRAL UNIVERSITY LIBRARY

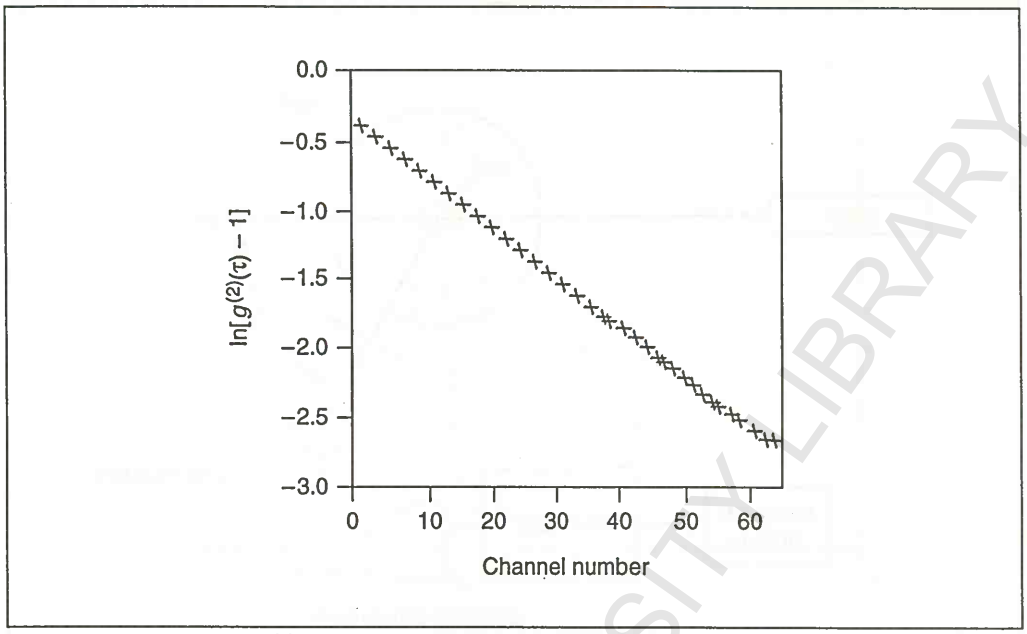


Figure 7.8.7 Plot of $\ln[g^{(2)}(\tau) - 1]$ versus channel number for the motility protein dynein. Channel number = delay time/sample time (Wells et al., 1990).

number of angles and at an additional extrapolation to zero angle (or Bragg vector q) is performed. The extrapolations to zero concentration and zero angle (or q) can be performed simultaneously on a biaxial extrapolation plot known as a dynamic Zimm plot (Burchard, 1992). If an angular extrapolation is not necessary, a scattering angle of 90° is usually chosen, and fixed-angle instruments are usually set at this angle (see Claes et al., 1992). At lower angles the problems due to supramolecular contamination are accentuated.

If a system is heterogeneous it is possible, at least in principle, to obtain a distribution of diffusion coefficients after various assumptions and mathematical manipulations of the autocorrelation data. The various methods of manipulation have been reviewed by Johnsen and Brown (1992) and several commercially available computer routines are available for performing these. A more simple way of representing heterogeneity is the polydispersity factor (PF) which is obtained by comparing linear with quadratic or quadratic with cubic-order fits of the normalized autocorrelation function decay data (Pusey, 1974).

Translational Diffusion Coefficient

The translational diffusion coefficient, $D_{20,w}$ or $D^0_{20,w}$, obtained by dynamic light scattering or boundary spreading in the ultracentrifuge, can be used to provide a number of useful characteristics about a biomolecular system.

Equivalent hydrodynamic radius (r_H)

The simplest deduction one can make from $D^0_{20,w}$ is the size of the biomolecule as represented by the equivalent hydrodynamic radius (also known as the Stokes radius)—r_H. What this means is, although the biomolecular system may not be a sphere at all, its diffusive behavior can be represented by an equivalent spherical particle of radius r_H. The value of r_H can be easily obtained from $D^0_{20,w}$ via the Stokes-Einstein relation:

$$r_H = \frac{k_B T}{6\pi\eta_{20,w} D^0_{20,w}}$$

Equation 7.8.7

where k_B, the Boltzmann constant, is equal to 1.379×10^{-16} erg/°K, T = 293.15°K and $\eta_{20,w}$ (the viscosity of water at 20.0°C) is equal to 0.01 poise. If D values are not corrected to standard conditions then the appropriate values for T and η have to be used.

Frictional coefficient (f)

The frictional coefficient provides a handle on molecular weight and shape as discussed below. Like r_H it can also be calculated simply from $D^0_{20,w}$ according to Equation 7.8.8.

$$f = \frac{RT}{N_A D^0_{20,w}}$$

Equation 7.8.8

BCU IAS CENTRAL UNIVERSITY LIBRARY

where T = 293.15°K and R = 8.314×10^{-7} erg/mol °K and N_A is Avogadro's number (6.022137×10^{23} mol).

Molecular weight

A more absolute description of biomolecular size than r_H is the molecular weight, *M*. Calculation of *M* from dynamic light scattering is less direct than from static measurements. However, combination of the $D^0_{20,w}$ value with the analogous parameter from sedimentation velocity experiments (*UNIT 7.5*) in the analytical ultracentrifuge—$s^0_{20,w}$ (in sec)—provides a popular route for obtaining the molecular weight of a biomolecule. The analogous equation to Equation 7.8.8 for the sedimentation coefficient is Equation 7.8.9.

$$f = \frac{M\left(1-\bar{v}\rho_{20,w}\right)}{N_A s^0_{20,w}}$$

Equation 7.8.9

Elimination of *f* between Equation 7.8.8 and Equation 7.8.9 yields the Svedberg (1927) equation (Equation 7.8.10):

$$M = \left(\frac{s^0_{20,w}}{D^0_{20,w}}\right)\left(\frac{RT}{1-\bar{v}\rho_{20,w}}\right)$$

Equation 7.8.10

where again T = 293.15°K. The variable $\bar{v}$ is the partial specific volume of the biomolecule (which can be measured from densimetry or from the composition of the biomolecule) and $\rho_{20,w}$ is the density of water at 20.0°C (0.9982 g/ml). It is possible in principle to measure both $D^0_{20,w}$ and $s^0_{20,w}$ simultaneously from analysis of the boundary shape in sedimentation velocity analytical ultracentrifugation, although in practical terms this needs data of very high quality and the absence of any effects of sample heterogeneity. Equation 7.8.10 also of course provides a route for measurement of $D^0_{20,w}$ if $s^0_{20,w}$ and *M* are known.

A simpler method for *M* measurement is possible if the conformation of the biomolecule is assumed and use is made of a power law relation, known as a Mark-Houwink relation (Equation 7.8.11):

$$D = KM^{-\varepsilon}$$

Equation 7.8.11

where $\varepsilon = 0.333$, 0.85, or 0.5 to 0.6 for sphere-, rod-, and coil-shaped molecules, respectively. From a calibration plot of log *D* versus log *M* for biomolecules of known *D* and *M* (Claes et al., 1992), *M* of an unknown can be found from its measured *D*. A conformation has of course to be assumed; the Equation 7.8.11 method for *M* evaluation is thus not as reliable as the method shown in Equation 7.8.10. Equation 7.8.11 can be taken even further in converting a distribution of diffusion coefficients into a molecular-weight distribution, although such distribution information is not as reliable (because of the approximations made) as that obtained with static light scattering methods coupled directly on-line to a gel-filtration column (Wyatt, 1992).

Conformation

The translational frictional coefficient (*f*) can be used directly to provide information on conformation. More convenient, however, is to use the corresponding dimensionless ratio called the frictional ratio (f/f_0), where f_0 is the frictional coefficient of a spherical particle of the same anhydrous mass and density as the biomolecule. The quantity f/f_0 can be calculated from $D^\circ_{20,w}$ by the relation in Equation 7.8.12.

$$f/f_0 = \left(\frac{k_B T}{6\pi\eta_{20,w}}\right)\left(\frac{4\pi N_A}{3\bar{v}M}\right)^{\frac{1}{3}}\left(\frac{1}{D^0_{20,w}}\right)$$

Equation 7.8.12

The frictional ratio depends intrinsically on the conformation, flexibility, and degree of solvent association (water plus other salt ions and any other solvent molecules) of the biomolecule. This degree of water association is termed the "hydration" of the biomolecule, δ, and is defined as the mass (in g) of associated solvent per g of anhydrous biomolecule. This associated solvent includes both chemically bound solvent and solvent physically entrained in the interstices of the molecule. The value of δ is typically between 0.2 and 0.5 g/g for proteins, although it is a notoriously difficult parameter to pin down with any accuracy.

The function defining the shape and flexibility of the biomolecule is the Perrin translational frictional function, *P*, illustrated by Equation 7.8.13:

$$P = (f/f_0)\left\{1+\frac{\delta}{\bar{v}\rho_0}\right\}^{-\frac{1}{3}}$$

Equation 7.8.13

ACU ... RAL UNIVERSITY LIBRARY

where ρ_0 is the density (g/ml) of the bound solvent. For a molecule that is fairly rigid on a time-averaged basis, the gross conformation can be specified using P in terms of the axial ratio of the equivalent hydrodynamic ellipsoid or in terms of sophisticated arrangements of spheres called hydrodynamic bead models. Computer programs are available for both types of modeling strategy (Harding et al., 1997; Garcia de la Torre et al., 1997)—although particularly with the latter, the diffusion coefficient should be used in conjunction with other hydrodynamic measurements to minimize uniqueness problems of the determined solution conformation.

Limitations of Dynamic Light Scattering Methods

As with static light scattering, the main limitations here relate to sample clarification. The technique is most suited to larger proteins and protein assemblies (M >100,000 Da) where problems due to contamination are less, and, as with static light scattering, the problems become disproportionally higher at the lower scattering angles. If the protein systems are approximately globular, then to a good approximation these low angles can be avoided and measurements at a single higher angle (usually 90°) can suffice. For nonspheroidal systems, angular extrapolations are necessary to eliminate rotational diffusion contributions. The user should also be wary of interpreting data from fixed-angle photometers, which do not permit such angular measurements.

The other limitations concern direct molecular-weight evaluation via Equation 7.8.11 and interpretation of the autocorrelation data in terms of polydispersity or size distribution. The user should be aware that such distributions are obtained purely by mathematical manipulation of the autocorrelation data, instead of involving a physical separation of different sizes (as with SEC-LALLS or SEC-MALLS), and the user is advised to be extremely cautious when interpreting such information produced by computer packages.

Samples for Analysis in Dynamic Light Scattering

Aqueous solvents should be of sufficient ionic strength to suppress charge effects. The loading concentrations required—which, as with static light scattering, should be measured *after* clarification—will depend principally on the size of the scatterer and the output from the laser. For example, if a 25-mW He-Ne laser is used, a loading concentration of at least ~1 mg/ml (and a volume of 2 to 3 ml) is required for a large protein assembly whose molecular weight (M) is ~5 × 10^6 Da. For proteins of molecular weight down to a lower limit of ~10,000 Da, more powerful lasers (~100 mW) and/or higher concentrations and/or longer experimental duration times are generally necessary to obtain meaningful results.

Temperature Control

D is very sensitive to temperature, mainly because of the dependence of diffusion on the viscosity of the solvent. Temperature needs to be controlled or, at the very least, monitored accurately during the measurement, and a water bath is highly desirable.

Clarification of Solutions and Scattering Cells

As with static light scattering, the scattering signal in dynamic light scattering is very sensitive to the presence of trace amounts of dust or supramolecular aggregation products; solutions and scattering vessels (called scattering cells or cuvettes) need to be scrupulously clean and free of particulates. Appropriate filtration, centrifugation of solutions, and washing of vessels with solvent is necessary. Special in-house cell-filling devices can be used along the lines described by Sanders and Cannell (1980); these can also be used for static light scattering.

Choice of Scattering Cuvette/Cell

The same criteria apply as with static light scattering (see Samples for Analysis in Static Light Scattering). The cell design depends on whether the photometer is a fixed-angle or multi-angle design.

Procedure for Making Scattering Measurements Using a Fixed-Angle Instrument with a Flow Cell

The following steps describe the injection procedure for a fixed (90°) angle photometer with a 20-mW infrared (780 nm) semiconductor laser (Claes et al., 1992).

1. The user should be satisfied that the protein/protein assembly is approximately globular. If not, then a multiangle instrument should be used (see Procedure for Making Scattering Measurements Using a Multiangle Instrument with a Conventional Cell).

2. Switch on the instrument and allow the laser to warm up 5 to 10 min.

BCU IAS CENTRAL UNIVERSITY LIBRARY

3. Inject pure ultrafiltered water or buffered solvent via a 0.1-μm filter to get the clean water count rate.

4. If the count rate is below the manufacturer's threshold, check the instrument alignment or increase the protein concentration.

5. Inject the protein solution in the same way as the water or buffer, using the appropriate filter, and check that the count rate meets the criteria specified by the manufacturer.

6. The choice of sample time and experimental duration time is normally done automatically.

7. Use the instrument's software to obtain the diffusion coefficient, and, where appropriate, the in-built calibration to directly obtain the approximate molecular weight.

8. Rinse and dry the flow cell of the photometer.

Procedure for Making Scattering Measurements Using a Multiangle Instrument with a Conventional Cell

1. Choose the appropriate cuvette or cell, using the same criteria (square versus cylindrical) as discussed above for static light scattering (see Samples for Analysis in Static Light Scattering).

2. Ensure that the cell is clear and free from dust and perform a check with pure ultrafiltered water, which should give only a negligibly small number of photon counts. Visually check for any "sparkling" in the laser beam passing through the water.

3. Remove the water from the cuvette, dry with ultrafiltered air, and inject the sample solution using the appropriate filter.

4. Set the goniometer for a scattering angle of 90°.

5. The choice of sample time and experimental duration time is normally done automatically.

6. Use the instrument's software to obtain the apparent diffusion coefficient, D_{app}. Inspect the autocorrelation decay plots (see Fig. 7.8.7). Nonlinearity can be due to sample polydispersity, particle asymmetry (effect of rotational diffusion phenomena), or particle settling (usually only observed for cellular systems, not for macromolecules). If it is not reasonable to assume an approximately globular system, then measure D_{app} at a series of angles and extrapolate to zero angle, but be aware of enhanced supramolecular contamination problems at the lower angles.

7. Remove the solution, rinse the cell with ultrafiltered pure water (after, if necessary, prior rinsing with detergent and mild acetic acid), dry with ultrafiltered air, and dry the cuvette.

Electrophoretic Light Scattering

Electrophoretic light scattering (ELS) uses dynamic light scattering to measure the velocity of migration, V, of a protein or other macromolecular system under the influence of an electric field, E (Langley, 1992; McNeil-Watson and Parker, 1992) The velocity V is related to the Doppler broadening, $\Delta\nu$, of the frequency of the incident laser radiation due to the velocity of the macromolecule. $\Delta\nu$ is related to V by Equation 7.8.14.

$$\Delta\nu = \left(\frac{2nV}{\lambda_0}\right)\sin\left(\frac{\theta}{2}\right)$$

Equation 7.8.14

The mobility, $U = V/E$, can thus be defined and hence the zeta potential, ζ:

$$U = \varepsilon\zeta/\eta$$

Equation 7.8.15

This procedure is however used more for the investigation of the stability of colloids rather than protein structure and so will not be discussed further here.

LITERATURE CITED

Bahls, D.M. and Bloomfield, V.A. 1977. Turbidimetric determination of bacteriophage molecular weights. *Biopolymers* 16:2797-2799.

Brown, W. (ed.) 1993. Dynamic Light Scattering: The Method and Some Applications. Clarendon Press, Oxford.

Burchard, W. 1992. Static and dynamic light scattering approaches to structure determination of biopolymers. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp. 3-22. Royal Society of Chemistry, London.

Claes, P., Dunford, M., Kenney, A., and Vardy, P. 1992. An on-line dynamic light scattering instrument for macromolecular characterisation. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp. 66-76. Royal Society of Chemistry, London.

Garcia de la Torre, J., Carrasco, B., and Harding, S.E. 1997. SOLPRO: Theory and computer program for the prediction of SOLution PROperties of rigid macromolecules and bioparticles. *Evr. Biophys. J.* 25:361-372.

Harding, S.E. 1997. Microbial laser light scattering. *Biotechnol. Genet. Eng. Rev.* 14:145-164.

BOSTON UNIVERSITY CENTRAL LIBRARY

Harding, S.E., Horton, J.C., and Colfen, H. 1997. The ELLIPS suite of macromolecular conformation algorithms. *Eur. Biophys. J.* 25:347-359.

Johnsen, R.M. and Brown, W. 1992. An overview of current methods of analysing QLS data. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp 77-91. Royal Society of Chemistry, London.

Johnson, P. 1993. Light scattering in the study of colloidal and macromolecular systems. *Int. Rev. Phys. Chem.* 12:61-87.

Johnson, P. and McKenzie, G.H. 1977. A laser light scattering study of haemoglobin systems. *Proc. R. Soc. Lond. B.* 199:263-278.

Johnson, P. and Ottewill, R.H. 1954. A light scattering study of diptheria toxin-antitoxin interaction. *Disc. Faraday Soc.* 18:327-337.

Jumel, K., Browne, P., and Kennedy, J.F. 1992. The use of low angle laser light scattering with gel permeation chromatography for the molecular weight determination of biomolecules. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp. 23-34. Royal Society of Chemistry, London.

Jumel, K., Fogg, F.J.J., Hutton, D.A., Pearson, J.P., Allen, A., and Harding, S.E. 1997. A polydisperse linear random coil model for the quaternary structure of pig colonic mucin. *Eur. Biophys. J.* 25:477-480.

Langley, K.H. 1992. Developments in electrophoretic laser light scattering and some biochemical applications. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.) pp. 151-160. Royal Society of Chemistry, London.

McNeil-Watson, F.K. and Parker, A. 1992. New tools for biochemists: Combined laser Doppler microelectrophoresis and photon correlation spectroscopy. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, and V.A. Bloomfield, eds.), pp. 59-65. Royal Society of Chemistry, London.

Pusey, P.N. 1974. Correlation and light beating spectroscopy *In* Photon Correlation and Light-Beating Spectroscopy (H.Z. Cummings and E.R. Pike, eds.) p. 387. Plenum, New York.

Sanders, A.H. and Cannell, D.S. 1980. Techniques for light scattering from hemoglobin. *In* Light Scattering in Liquids and Macromolecular Solutions (V. Degiorgio, M. Corti, and M. Giglio, eds.) pp. 173-182. Plenum, New York.

Stacey, K.A. 1956. Light Scattering in Physical Chemistry. Butterworths, London.

Svedberg, T. 1927. Nobelvortrag gehalten zu Stockholm am 19. Mai 1927. *Kolloid-chem. Beih.* 26:230-244.

Tanford, C. 1961. Physical Chemistry of Macromolecules. Chapters 4 and 5. John Wiley & Sons, New York.

Wells, C., Molina-Garcia, A.D., Harding, S.E., and Rowe, A.J. 1990. Self-interaction of dynein from *Tetrahyamena* cilia. *J. Muscle Res. Cell Motil.* 11:344-350.

Wyatt, P.J. 1992. Combined differential light scattering with various liquid chromatography separation techniques. *In* Laser Light Scattering in Biochemistry (S.E. Harding, D.B. Sattelle, V.A. Bloomfield, V.A. eds.) pp. 35-58. Royal Society of Chemistry, London.

Yau, W.W., Kirkland, J.J., and Bly, D.D. 1979. Modern Size Exclusion Chromatography. John Wiley & Sons, New York.

Zimm, B.H. 1948. *J. Chem. Phys.* 16:1093-1099.

KEY REFERENCES

Harding S.E., Sattelle, D.B. and Bloomfield, V.A. (eds.) 1992. Laser Light Scattering in Biochemistry. Royal Society of Chemistry, London.

Huglin, M.B. (ed.) 1972. Light Scattering from Polymer Solutions. Academic Press, London.

Contributed by Stephen E. Harding and Kornelia Jumel
University of Nottingham School of Biology
Sutton Bonington, United Kingdom.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Measuring Protein Thermostability by Differential Scanning Calorimetry

Every chemical reaction is accompanied by heat effects that, at constant pressure, are described by the enthalpy of the process. Protein folding/unfolding, being an isomerization reaction between different macroscopic states, is no exception (Sturtevant, 1972; Privalov and Khechinashvili, 1974; Krishinan and Brandts, 1978). Temperature-induced changes in the enthalpy of the macroscopic states of a protein are of particular significance. Temperature and enthalpy are coupled extensive and intensive thermodynamic parameters. Thus, the functional relationship between enthalpy and temperature, $f = H(T)$, includes all thermodynamic information about the macroscopic states in the system. This information can be extracted using the formalism of equilibrium thermodynamics.

The $H(T)$ function of a system can be determined experimentally from calorimetric measurements of its heat capacity at constant pressure $C_p(T)$, over a broad temperature range:

$$C_p(T) = \left(\frac{\partial H}{\partial T}\right)_p \text{ and thus } H(T) = \int_{T_0}^{T} C_p(T)dT + H(T_0)$$

Equation 7.9.1

The heat capacity function, $C_p(T)$, of a system can be directly measured by differential scanning calorimetry (DSC), making DSC the method of choice to study the conformational stability of biological macromolecules and proteins in particular.

This unit presents step-by-step protocols for DSC, including sample preparation and interpretation of the results in the simplest cases (see Basic Protocol) as well as calibration of the apparatus (see Support Protocols 1, 2, and 3) and maintenance of DSC cells (see Support Protocol 4). Various general experimental considerations, including possible pitfalls and errors, are discussed in the Commentary (see Critical Parameters; see Troubleshooting).

STRATEGIC PLANNING

To study biological macromolecules using DSC instrumentation, a number of requirements must be satisfied: the measurements must be performed (1) in a physiological milieu, i.e., aqueous solution; (2) on protein samples that are dilute—i.e., at concentrations as low as 1 mg/ml (0.1%)—to minimize intramolecular interactions; and (3) in a small volume, since isolating and purifying protein samples is costly (in most cases only a few milligrams of protein are available, which at 1 mg/ml is only a few milliliters). Under such conditions the partial heat capacity of a protein in aqueous solution is on the order of 0.03% to 0.5%, and this is measured on the background of >99.5% heat capacity of solvent. Such measurements require extreme sensitivity, reproducibility, and signal stability from the DSC instrumentation. This implies that particular care should be taken with the experimental procedures for sample preparation and routine DSC operation. Currently two companies manufacture DSC instruments capable of working with solutions of biological macromolecules (Calorimetric Science and Microcal; see *SUPPLIERS APPENDIX*). These instruments are fully automated for control and for data collection, handling, and analysis using PC computers. The general principles of DSC instrument operation are outlined by Privalov et al. (1995) and Plotnikov et al. (1997).

Contributed by George I. Makhatadze
Current Protocols in Protein Science (1998) 7.9.1-7.9.14
Copyright © 1998 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

In performing a DSC experiment, it is important to keep in mind that DSC operates in differential mode. This means that the heat capacity of the protein in aqueous solution is measured relative to the heat capacity of the surrounding buffer. In an ideal situation, when the heat capacities and volumes of both sample and reference cells are precisely the same, a single protein/buffer scan will suffice. However, in reality, sample and reference cells are not identical and this difference has to be taken into consideration. Experimentally this requires the use of an internal instrument standard—the buffer/buffer scan, or baseline of the instrument. Before starting the protein scan it is very important to establish the stability of the baseline, i.e., its relative position and shape. Most contemporary calorimeters have very high baseline stability and reproducibility. Before starting to prepare the protein it is always a good idea to take several overnight buffer/buffer baselines and compare these to make sure that they are reproducible in accordance with previous observations. If this criterion is satisfied, one may start preparing the protein sample (see Critical Parameters for crucial steps in sample preparation).

When setting the DSC control program, one of the most important parameters is the heating rate. Several considerations have to be taken into account. First, the higher the scan rate, the higher the sensitivity of the instrument. The increase in sensitivity is linear with respect to the heating rate: i.e., sensitivity at a heating rate of 2°/min is twice as high as at 1°/min. One needs to keep in mind that the increase in sensitivity does not mean an increase in the signal-to-noise ratio. Second, if the expected transition is very sharp—i.e., occurs within a few degrees—a high heating rate will affect the shape of the heat absorption profile and lead to an error in the determination of all thermodynamic parameters for this transition. This error will be particularly large for the transition temperature. Third, the higher the heating rate, the less time the system has to equilibrate. So for slow unfolding/refolding processes it is advisable to use low heating rates. In practice, for small globular proteins exhibiting reversible temperature-induced unfolding, the maximal heating rate (on most DSC instruments) of 2°/min is acceptable. For large proteins lower heating rates (0.5° to 1°/min) are more appropriate. In the case of fibrillar proteins, which usually exhibit very narrow transitions, a heating rate of 0.1° to 0.25°/min is most suitable.

BASIC PROTOCOL

DIFFERENTIAL SCANNING CALORIMETRY

Before beginning calorimetry, an instrument baseline is established by a buffer/buffer scan. If there are changes in the shape and/or relative position of the buffer/buffer scan, the cells are probably dirty (see Support Protocol 4 for cleaning procedure).

Materials

Protein solution, dialyzed (*UNIT 4.4 & APPENDIX 3.B*) and in dialysis bag (see Critical Parameters)
Equilibration buffer (buffer solution used for the last dialysis; see Critical Parameters for discussion of buffer selection)
Pasteur pipet or automatic pipettor
Spectrophotometer and quartz cuvettes of appropriate optical path length (depending on the expected protein concentration)
2-ml all-glass syringe (e.g., Becton Dickinson 2-cc Yale syringe)
Precut needle supplied with the DSC instrument
Differential scanning calorimeter (DSC), calibrated (see Support Protocols 1, 2, and 3)

PAU IASI CENTRAL UNIVERSITY LIBRARY

1. Turn on the spectrophotometer.

   This step is required only if a spectrophotometric method of determining protein concentration is used. See Critical Parameters for discussion of appropriate methods for determining protein concentration, which is crucial to obtaining accurate DSC data.

2. Transfer the protein solution from the dialysis bag into a 1.5-ml microcentrifuge tube using a clean Pasteur pipet or an automatic pipettor, and centrifuge 15 to 20 min at ~12,000 to 13,000 × *g*, 4°C.

   This step removes insoluble material.

3. While the protein is centrifuging, fill two quartz spectrophotometer cuvettes with equilibration buffer and take a buffer/buffer spectrum.

   This step establishes the buffer-buffer baseline of the spectrophotometer, which is particularly important with buffers containing absorbing components (e.g., dithiothreitol).

4. Empty both calorimetric cells, and rinse an all-glass syringe and attached needle (precut needle supplied with the DSC instrument) at least five to seven times with equilibration buffer.

5. Rinse one calorimetric cell (the reference cell) as follows: fill up the syringe, pump out all air bubbles, and insert the needle into the cell, letting the plunger slide down under its own weight (needles are precut to a specific length so that the tip of the needle will be barely above the bottom of the cell). Once the buffer appears in the overflow reservoir, start abruptly pumping buffer in and out of the cell, for a total of at least eleven to thirteen syringe volumes. Then fill the cell with more buffer.

   The abrupt pumping will force trapped air bubbles out of the cell.

   Calorimetric cells can be refilled using disposable plastic syringes as well (e.g., in the case of Ca^{2+}-binding proteins, where it is desirable to avoid glass syringes since they will "leak" calcium). In this case the operator should slowly inject solution into the cell until the solution appears in the overflow reservoir, and then briskly pump the solution in and out of the cell. During the last step, the solvent level should never get below the level of the overflow reservoir.

6. Repeat rinsing procedure with the second calorimetric cell (sample cell) and leave it empty. Remove all buffer leftovers from the overflow reservoir.

7. Stop the centrifuge. Using the syringe, transfer the supernatant of the protein solution into a spectrophotometric cuvette and record the protein spectrum.

   It is advisable not to try and remove all of the supernatant, but to leave last 50 to 70 µl in the microcentrifuge tube.

8. Transfer the protein solution from the spectrophotometric cuvette into the sample calorimetric cell.

   Follow the same technique as for filling the reference cell (see step 5).

9. Apply pressure, maintaining at least 1.5 atm (~22 psi) in the cell.

   All calorimetric experiments should be performed under excess pressure (see Critical Parameters).

10. Select scan parameters and start the DSC control program.

    See Strategic Planning for discussion of control parameters; see Anticipated Results for information on data analysis procedures.

BOUVIER / CENTRAL UNIVERSITY LIBRARY

CALIBRATION OF THE DSC INSTRUMENT

DSC instruments are precalibrated by their manufacturers for temperature scale and heat capacity. It is unlikely that the characteristics of an instrument will change, but it is nonetheless advisable to check them once or twice a year. Two other calibrations should be done in the lab: electrical calibration (y axis of the instrument readout) using the system built into the instrument (see Support Protocol 1), and temperature calibration (x axis) using either hydrocarbon capsules obtained from the manufacturer (see Support Protocol 2) or lipid suspensions that can be prepared in the laboratory (see Support Protocol 3).

SUPPORT PROTOCOL 1

Electrical Calibration

All DSC instruments have a built-in capability to perform heat capacity calibration using an electrical pulse. Electrical calibration is performed using a special built-in circuit with a precisely measured resistance. A precise electrical current applied to this circuit generates a defined amount of power, δW, supplied to one of the cells. Exactly the same amount of power will be supplied to the other cell by the cell heaters in order to compensate for the inequality of temperature between the sample and reference cells. The power generated by the heaters of the instrument is related to the apparent heat capacity, δC_p; the y-axis value generated by the instrument is $\delta C_p = \delta W/(dT/dt)$, where dT/dt is the heating rate. Follow the manufacturer's suggested procedures for electrical calibration.

SUPPORT PROTOCOL 2

Temperature Calibration Using Hydrocarbon Capsules

Temperature calibration is performed by melting pure hydrocarbons of known melting temperatures. The instrument manufacturers supply capsules containing hydrocarbons that can be inserted into a DSC cell filled with water (follow the manufacturer's suggested procedure). Alternatively, aqueous suspensions of certain lipids with known melting temperatures can be used (see Support Protocol 3).

SUPPORT PROTOCOL 3

Temperature Calibration Using Lipid Suspensions

The following procedure for preparing lipid suspensions was suggested by Mabrey and Sturtevant (1976).

Additional Materials (also see Basic Protocol)

Lipid: dipalmitoylphosphatidylcholine (DPPC) or distearoylphosphatidylcholine (DSPC)
10 mM sodium phosphate buffer, pH 7.0 (*APPENDIX 2E*)

Vortex mixer
Water bath, 60°C

1. Add the lipid (DPPC or DSPC) to 10 mM sodium phosphate buffer, pH 7.0, to a 0.5 to 1.0 mg/ml final lipid concentration.
2. Heat the suspension at 60°C for several minutes, then vortex vigorously.
3. Fill the reference calorimetric cell with 10 mM sodium phosphate buffer; fill the sample cell with the lipid suspension.
4. Set the heating rate of the calorimeter at 0.1° to 0.25°C/min.

If it is possible to set the response time of the instrument, use the fastest available.

P.G.U. IASI / CENTRAL UNIVERSITY LIBRARY

5. Start the scan.

The DPPC suspension has two transitions, a low-temperature, broad transition at 35.3°C and a high-temperature, sharp transition at 41.4°C. For the DSPC suspension the low-temperature transition is at 51.5°C and the high-temperature transition is at 54.9°C.

MAINTENANCE AND CLEANING OF DSC CELLS

SUPPORT PROTOCOL 4

DSC cells that become clogged by protein aggregates must be cleaned carefully in order to provide a reproducible baseline and thus accurate results. Proper use of calorimeters (i.e., avoidance of aggregating systems) minimizes the need for this sometimes very tedious procedure. Depending on the composition of the calorimetric cells (gold, platinum, titanium, tungsten), the chemical resistance of the material and the degree of contamination, a number of different procedures can be recommended.

Additional Materials (also see Basic Protocol)

1 M HCl *or* 1 M NaOH
Distilled water
Concentrated formic acid
95% (190-proof) ethanol
2% (w/v) SDS, or other detergent
Concentrated nitric acid

1. Rinse calorimetric sample cell with one syringe volume of 1 M HCl or 1 M NaOH. Rinse the cells well with water (~20 to 30 syringes-full; see Basic Protocol, step 5). Fill both cells and record a water-water scan. If the baseline does not return to its original position, proceed with step 2.

   The choice between acid and base depends on the properties of the protein aggregates formed. This can be checked by titrating the aggregated protein sample removed from the calorimeter with acid or base (half and half).

   IMPORTANT NOTE: *Do not use NaOH with cells made out of tungsten or titanium, as they will corrode.*

2. Fill the sample cell with formic acid and let it sit for ~10 to 15 min at room temperature. Rinse the cells well with water (~20 to 30 syringes-full), fill both cells, and record a water-water scan. If the baseline does not return to its original position, proceed with step 3.

3. Fill both cells with formic acid and scan several times up to 70°C. Rinse the cells well with water (~20 to 30 syringes-full), fill both cells, and record a water-water scan. If the baseline does not return to its original position, proceed with step 4.

4. Fill both cells with 95% ethanol and scan several times to 70°C. Rinse the cells well (~20 to 30 syringes-full of water), fill both cells, and record a water-water scan. If the baseline does not return to its original position, proceed with step 5.

5. Fill both cells with detergent and scan several times to the maximal temperature allowed by the instrument. Rinse very well (it is quite difficult to rid cells of detergent, but ~40 to 50 syringes-full of water per cell usually suffice). Fill both cells and record a water-water scan. If the baseline does not return to its original position, proceed with step 6.

6. Fill both cells with nitric acid and scan to 70°C several times. Rinse well (~20 to 30 syringes-full), fill both cells with water, and record a water-water scan.

   Particular care should be taken to prevent exposing the pan or pressure sensor (in some designs) to the nitric acid fumes, which will cause rusting. Put wet cotton wool or a Kimwipe on top of the cell outlets and pressure-sensor hole.

BOSTON BIOMEDICAL RESEARCH INSTITUTE CENTRAL UNIVERSITY LIBRARY

If none of these treatments helps, a possible alternative reason for changes in the shape and position of the baseline is changes in the intrinsic thermal properties of the cells that can occur over time. Contact the instrument manufacturer for advice.

COMMENTARY

Background Information

The major advantage of using DSC to study protein stability compared to other methods for monitoring temperature-induced unfolding of proteins is that it provides direct information about the modes of unfolding. This technique is unique for establishing whether the unfolding process is a two-state or a multi-state process. It also provides direct measurement of the energetics of the unfolding transition. Another important parameter directly measured by DSC is the heat capacity change upon unfolding, which can be used to compute the temperature dependencies of enthalpy, entropy, and Gibbs energy functions. Until recently, the major drawback to DSC was the requirement for relatively large amounts of a protein sample. Contemporary DSC instruments have overcome this and allow measurements at spectroscopic concentrations (0.1 mg/ml).

Critical Parameters

Protein purity

The purity of the sample should be maintained very scrupulously. Traditional electrophoresis (native, *UNIT 10.1*, and SDS-PAGE, *UNIT 10.3*) as well as emerging, more sensitive methods (MALDI-TOF or ion-spray mass spectroscopies, *UNIT 16.2*) should be used to check sample purity. Impurities can affect both the shape of the calorimetric profiles and the estimation of the thermodynamic parameters of unfolding through errors in determination of the sample concentration (see below).

Buffer solutions

Careful attention should be paid to the buffer composition. Two factors are particularly important.

1. Since calorimetric experiments are performed over a broad temperature range, the pH of the buffer should have a low temperature dependence. For this reason the very popular "physiological" buffer Tris is not the best choice. Glycine, sodium or potassium acetate, sodium or potassium phosphate, sodium cacodylate, and sodium citrate are all suitable buffers. See Izatt and Christensen (1976) for tabulated data on the physicochemical properties of various buffers.

2. Protein unfolding is usually accompanied by changes in the ionization state of certain moieties within the protein (i.e., amino acid functional groups). The heat of ionization can be very large (e.g., 30 to 44 kJ/mol for histidine, depending on pH) and can obscure the "true" conformational enthalpy of unfolding. In some cases this problem can be avoided by selecting buffers with heats of ionization close to the heats of ionization of the protein groups. In this case the heat of proton release/absorption for the protein will be compensated by the heat of proton absorption/release for the buffer. This can be achieved in the acidic pH ranges of 2.0 to 3.2 (with the glycine/HCl buffer system) and 3.5 to 5.0 (with the sodium acetate/acetic acid buffer system). In this overall pH range (2.0 to 5.0), the carboxyl groups of aspartic and glutamic acid are titrated, and their heat of ionization will be counterbalanced by the heat of ionization of the carboxyl groups of the buffer. In neutral and alkaline pH ranges, such compensation is not possible and corrections for the effects of functional group ionization should be made during data analysis (see Anticipated Results).

Equilibration of the protein solution with the solvent

Equilibration of the protein with buffer is a vital part of sample preparation for calorimetric experiments. One should keep in mind that what the DSC instrument measures is the difference in heat capacity between buffer and protein solutions. This means that any small variations in buffer composition will result in an incorrect absolute value and an incorrect slope of the heat capacity profile. Direct dissolution of a protein sample in the corresponding buffer gives an equimolar concentration of the low-molecular-weight component (the buffer). In the case of ionic compounds, this will lead to nonidentical buffer compositions, since the ionic solute—i.e., protein—will directly interact with the buffer components, thereby decreasing its bulk concentration.

The best way to achieve equilibrium between buffer and protein solution is dialysis. The protein sample should be dialyzed against at least two changes of buffer (see *UNIT 4.4*). The appropriate dialysis time will vary depending

PCU IASI CENTRAL UNIVERSITY LIBRARY

on the molecular weight cutoff (MWCO) of the dialysis membrane: the lower the MWCO, the longer the dialysis should be. For membranes with MWCO of 3500, 6 hr per change are enough. Higher-MWCO membranes equilibrate somewhat faster. It is advisable to filter the buffer solutions prior to dialysis using a 0.45-μm filter to eliminate all insoluble material.

Clarification of the solution

After dialysis, all insoluble particles should be carefully removed from the protein solution. This can be achieved either by passing the protein solution through a 0.22- to 0.45-μm filter or by microcentrifuging the solution for 15 to 20 min at ~12,000 × *g*. Centrifugation is preferable for several reasons: (1) quite often there is nonspecific adsorption of protein to the filtration membrane; (2) the filtration membrane can contain compounds that will be washed off into the protein solution, thus changing the buffer composition of the protein solution compared to the equilibration buffer; and (3) filtration membranes are costly.

Concentration of the protein solution

Knowledge of the exact protein concentration *after* dialysis against the corresponding buffer is another important prerequisite to DSC. Among all available methods for determining protein concentration, the spectrophotometric method seems to be most accurate and rapid. This method requires (1) a standard double-beam UV/visible spectrophotometer with resolution of at least 0.002 absorption units, and (2) knowledge of the extinction coefficient of the protein at a given wavelength. The extinction coefficient can be determined experimentally by determination of the nitrogen content (Jaenicke, 1974) or by quantitative amino acid analysis (*UNIT 3.2*). Alternatively, if the protein of interest contains at least one tyrosine or tryptophan residue or disulfide bond, the extinction coefficient can be calculated from the number of aromatic residues and disulfide bonds using an empirical equation (Gill and von Hippel, 1989):

$$\varepsilon_{280\ nm}^{0.1\%,1\ cm} = \frac{5690N_{Trp} + 1280N_{Tyr} + 120N_{SS}}{MW}$$

Equation 7.9.2

where N_{Trp}, N_{Tyr}, and N_{SS} are the number of tryptophans, tyrosines, and disulfide bonds, respectively, and MW is the molecular weight of the protein in daltons. A more elaborate procedure, described by Pace et al. (1995), improves the accuracy of this calculation.

For proteins that do not contain any aromatic residues, the method of Scopes (1974), which also uses spectrophotometry, gives fast and accurate results (see *UNIT 3.1*).

Light scattering in spectrophotometric measurements of protein concentration can be a significant problem. This can be taken into account using a procedure suggested by Winder and Gent (1971).

Pressure in the calorimetry cell

All calorimetric experiments are performed under excess pressure. There are two main reasons for this. (1) The solubility of gases in aqueous solution is temperature dependent—thus, the lower the temperature, the higher the solubility. Excess pressure prevents air bubble formation in the cell at elevated temperatures. In some cases this can also be achieved by degassing protein solutions by exposing them to vacuum. However, degassing inevitably leads to a change in the solvent composition due to evaporation, which will in turn introduce a distortion in the calorimetric profile. (2) Excess pressure keeps sample evaporation from the calorimetric cell to a minimum. It is important to obtain heat capacity measurements for as broad temperature range as possible; for some proteins, the unfolding transition is not complete until the temperature is above the boiling point of water. Excess pressure will elevate the boiling temperature, and allow measurements at temperatures above the boiling point of water (100°C) in these cases. It is recommended that the pressure in the cell be kept to at least 1.5 atm (~22 psi) in all experiments.

Troubleshooting

The most frequent problem in DSC experiments is irreversibility of the unfolding transition. This applies particularly to large proteins that tend to aggregate after unfolding. The sensitivity of contemporary DSC instruments provides reliable heat capacity profiles at concentrations as low as 0.1 mg/ml for proteins 6 to 17 kDa in size, and the concentration can be even lower for larger proteins. Low protein concentration decreases the probability of aggregation. However, even at low protein concentrations, finding the proper solvent conditions to prevent aggregation is not a simple task. A rule of thumb is that at a pH value close to the isoelectric point of the protein, aggregation

PCU INSTITUTIONAL UNIVERSITY LIBRARY

after unfolding is highly probable. High concentrations of neutral salts also promote aggregation. In many cases, aggregation can be avoided by keeping the ionic strength of the solvent very low (10 to 50 mM) and by using pH values a couple of units away from (usually below) the isoelectric point.

Anticipated Results

Calculation of the partial heat capacity profile

The first step in calculating the partial heat capacity profile is subtraction of the scan rate–normalized baseline (buffer/buffer scan) from the scan rate–normalized protein scan (protein/buffer) to obtain $\Delta C_p^{app}(T)$, the heat capacity difference between sample and reference cells at temperature T. Knowledge of $\Delta C_p^{app}(T)$ allows calculation of the partial heat capacity of the protein at temperature T, $C_{p,pr}^{exp}(T)$, as:

$$C_{p,pr}^{exp}(T) = C_p,buf(T)\cdot\frac{\overline{V}_{pr}(T)}{\overline{V}_{buf}(T)} - \frac{\Delta C_p^{app}(T)}{m_{pr}(T)}$$

Equation 7.9.3

where $C_{p,buf}(T)$ is the heat capacity of buffer at temperature T, $\overline{V}_{pr}(T)$ is the partial volume of the protein at temperature T, $\overline{V}_{buf}(T)$ is the partial volume of buffer at temperature T, and $m_{pr}(T)$ is the mass of the protein in the calorimetric cell at temperature T. The mass of the protein in the cell is calculated as $m_{pr}(T) = v_{cell}(T)\cdot W_{pr}(T)$, where $v_{cell}(T)$ is the volume of the cell at temperature T and $w_{pr}(T)$ is the concentration of the protein at temperature T.

In practice, however, for aqueous buffers

$$\frac{C_{p,buf}}{V_{buf}} \cong \frac{C_{p,H_2O}}{V_{H_2O}}$$

Equation 7.9.4

and $m_{pr}(T) \cong$ constant, because

$$\frac{\partial v_{cell}(T)}{\partial T} \cong \frac{\partial w_{pr}(T)}{\partial T}$$

Equation 7.9.5

so that Equation 7.9.3 is simplified to:

$$C_{p,pr}^{exp}(T) = \frac{C_{p,H_2O}}{\overline{V}_{H_2O}}\cdot\overline{V}_{pr} - \frac{\Delta C_p^{app}(T)}{m_{pr}}$$

Equation 7.9.6

where only one term, $\Delta C_p^{app}(T)$, is temperature dependent. In any case, both equations require knowledge of the partial volume of the protein, $\overline{V}_{pr}$. This parameter can be calculated from the amino acid composition of the protein using an additivity scheme described by Makhatadze et al. (1990). Calculations can also be performed via the World Wide Web at the URL *http://www.ttu.edu/~chem/makhatadze/Vpartial.html.*

Interpretation of DSC profiles

Two important pieces of information should be established prior to the analysis of the calorimetric profiles. These will help to decide whether the formalism of equilibrium thermodynamics is applicable to the protein being studied.

1. Reversibility of the unfolding transition should be examined under several conditions by rescanning the sample. Reversibility depends on the upper temperature of heating. The unfolding transition can be considered reversible if rescanning after the sample has been heated to the temperature just past the completion of the transition leads to the recovery of at least 85% to 90% of the initial endotherm. If no endotherm is observed after rescanning, the unfolding reaction is irreversible. Thermodynamic analysis for irreversible transitions is possible only in special cases (Sanchez-Ruiz, 1992).

2. The equilibrium conditions of reversible unfolding should be examined to find the optimal scan rate that permits establishment of equilibrium between native and unfolded states at all temperatures—i.e., the scan rate should not be faster than the folding/unfolding rates. This can be accomplished by comparing the heat capacity profiles obtained on the same exact sample at two different heating rates, usually 2°/min and 0.5°/min. If the difference in the transition temperature exceeds 0.5° to 0.7°C, additional experiments should be performed at a lower heating rate (i.e., 0.1°/min) and compared to the 0.5°/min experiment. If there is still a difference in the transition temperature, all experiments at a given solvent composition should be performed at several different heating rates and the transition temperatures should be extrapolated to a zero heating rate (for more details see Yu et al., 1994).

NOTE: Sensitivity of DSC is reversibly proportional to the heating rate, so these experiments should be performed at protein concentrations high enough to obtain the transition temperature with a relatively high accuracy.

ASU IASI CENTRAL UNIVERSITY LIBRARY

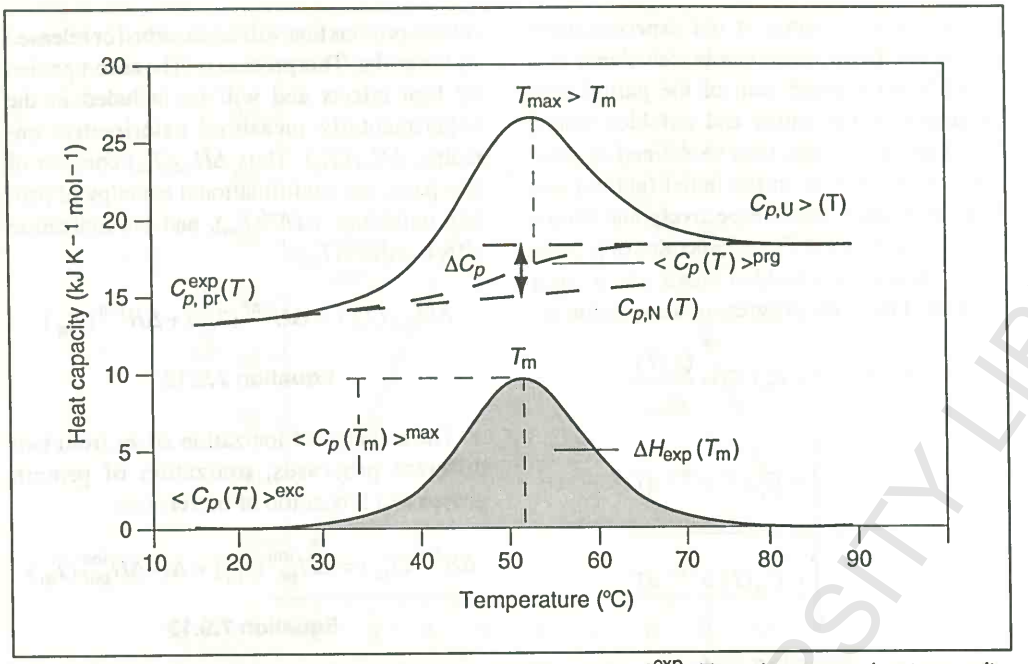


Figure 7.9.1 The relation of the partial molar heat capacity, $C_{p,pr}^{exp}(T)$, to the excess heat capacity, $<C_p(T)>^{exc}$, progress heat capacity, $C_p^{prg}(T)$, and partial heat capacities of the native, $C_{p,N}(T)$, and unfolded, $C_{p,U}(T)$, states.

By visual inspection of the calorimetric profiles one can establish whether the unfolding transition represents a single endotherm (as shown in Fig. 7.9.1) or is more complex. These two cases are analyzed differently.

If the DSC profile is a single endotherm

Heat capacity functions for the native and unfolded states. The partial specific heat capacity of the native state, $C_{p,N}$, ranges from 1.25 to 1.80 J/(K·g) at 25°C depending on the protein (Makhatadze, 1998). $C_{p,N}$ has a linear dependence on temperature with a slope from 0.005 to 0.008 J/(K^{-2} g), also depending on the protein (Makhatadze, 1998). The partial specific heat capacity of the unfolded state, $C_{p,U}$, in the temperature range from 5° to 100°C is always higher than the heat capacity of the native state. At 25°C, $C_{p,U}$ values for different proteins range from 1.85 to 2.2 J/(K·g), whereas at 100°C these values are higher, from 2.1 to 2.4 J/(K·g) (Makhatadze, 1998). The partial heat capacity of the unfolded state has a nonlinear dependence on temperature (e.g., Wintrode et al., 1994). It increases gradually (with a slope comparable to that for the native state) and approaches a constant value at 60° to 75°C. At temperatures above 60° to 75°C, $C_{p,U}$ does not depend on temperature. As a result of the different temperature dependencies of the heat capacities of the native and unfolded states, the heat capacity change upon protein unfolding, $\Delta C_p = C_{p,U} - C_{p,N}$, appears to be a temperature-dependent function.

Excess heat capacity profile. The so-called excess heat capacity function $<C_p(T)>^{exc}$ is the heat capacity caused exclusively by the heat absorbed due to the unfolding reaction. $<C_p(T)>^{exc}$ thus is expressed as a difference between the experimental, $C_p^{exp}(T)$, and progress, $C_p^{prg}(T)$, heat capacity functions of a protein:

$$< C_p(T) >^{exc} = C_p^{exp}(T) - C_p^{prg}(T)$$

Equation 7.9.7

where the progress heat capacity is defined by the intrinsic (partial) heat capacity of the protein in the entire temperature range:

$$C_p^{prg}(T) = F_N(T) \cdot C_{p,N}(T) + F_U(T) \cdot C_{p,U}(T)$$

Equation 7.9.8

where $C_{p,N}(T)$ and $C_{p,N}(T)$ are the temperature dependencies of the partial heat capacities of the native and unfolded states, respectively. Because of the difference in the partial heat capacities of the native and unfolded states, in the temperature range where transition occurs,

B.C.U. „M. Eminescu" IAŞI
CHIMIE

BUCHAREST UNIVERSITY LIBRARY

progress heat capacity is not experimentally accessible. However, it can be calculated as a population-weighted sum of the partial heat capacities of the native and unfolded states. $F_N(T)$ and $F_U(T)$ can thus be defined as fractions of the protein in the initial (native) and final (unfolded) states, respectively, and consequently $F_N(T) + F_U(T) = 1$. Fractions of protein in the native or unfolded states are in turn calculated from the progress of the reaction as:

$$F_U(T) = 1 - F_N(T) = \frac{Q_i(T)}{Q_{tot}} = \frac{\int_0^T < C_p(T) >^{exc} dT}{\int_0^\infty < C_p(T) >^{exc} dT}$$

Equation 7.9.9

where Q_{tot} is the total heat absorbed upon unfolding and $Q_i(T)$ is the heat absorbed at temperature T.

Transition temperature, T_m. The transition temperature, T_m, is defined as the temperature at which the areas under the excess heat capacity profile on both sides of T_m are equal, i.e.:

$$\int_0^{T_m} < C_p(T) >^{exc} dT = \int_{T_m}^\infty < C_p(T) >^{exc} dT$$

Figure 7.9.10

T_m always differs from the temperature of maximum partial heat capacity, T_{max}, due to the difference in the heat capacities between the native and unfolded states, $\Delta C_p = C_{p,U} - C_{p,N}$. $T_{max} - T_m$ decreases with an increase in the relative ratio of the area under the excess heat capacity profile to ΔC_p.

Calorimetric enthalpy, ΔH_{exp} (T_m). Calorimetric enthalpy is defined as the area under the excess heat capacity function, i.e.,

$$\Delta H_{exp}(T_m) = Q_{tot} = \int_0^\infty < C_p(T) >^{exc} dT$$

Equation 7.9.11

and is referred to as the enthalpy at the transition temperature.

Enthalpy of the conformational transition, $\Delta H^{cnf}(T_m)$. When a protein unfolds, internal groups buried in the native state can bind or release protons that will be absorbed or released by the buffer. This process will be accompanied by heat effects and will be included in the experimentally measured calorimetric enthalpy, $\Delta H_{cal}(T_m)$. Thus $\Delta H_{cal}(T_m)$ consists of two parts, the conformational enthalpy of protein unfolding, $\Delta H^{cnf}(T_m)$, and the ionization effects, $\Delta H^{ion}(T_m)$:

$$\Delta H_{exp}(T_m) = \Delta H^{cnf}(T_m) + \Delta H^{ion}(T_m)$$

Equation 7.9.12

The enthalpy of ionization arises from two different processes, ionization of protein groups and ionization of buffer; i.e.:

$$\Delta H^{ion}(T_m) = \Delta H^{ion}_{pr}(T_m) + \Delta\nu \cdot \Delta H^{ion}_{buf}(T_m)$$

Equation 7.9.13

where $H^{ion}_{pr}(T_m)$ is the enthalpy of ionization of protein, $H^{ion}_{buf}(T_m)$ is the enthalpy of ionization of buffer, and ν is the change in the protonation state of the protein (see below). When the enthalpies of ionization of protein groups and of the buffer are close, the calorimetric and conformational enthalpies are also close to each other, i.e, $\Delta H_{cal}(T_m) \cong \Delta H^{cnf}(T_m)$. Such a situation occurs in glycine or acetate buffer systems with an acidic pH. In the acidic pH range, the most probable titratable groups of the protein are the carboxylic acids (aspartic and glutamic acid). Their heats of ionization are very close to the heats of ionization for carboxylic groups of glycine or acetic acid. However, at neutral or alkaline pH, titratable groups (histidine, lysine, arginine, tyrosine, and cysteine) have heats of ionization very different from those of the buffer systems used (e.g., phosphate and cacodylate). In these cases, differences between calorimetric and conformational enthalpies can be significant (see, for example, Yu et al., 1994).

Changes in the protonation state of the protein upon unfolding, $\Delta\nu$. Under certain solvent conditions protein unfolding is accompanied by a change in the protonation state of the protein, $\Delta\nu$. This difference in the number of protons bound to the unfolded, ν_U, and native, ν_N, states can be determined from the dependence of the transition temperature on the pH values of the solution dT_m/dpH as:

$$\Delta\nu = \nu_U - \nu_N = \frac{\Delta H_{exp}(T_m)}{2.303 \cdot R \cdot T_m^2} \cdot \frac{dT_m}{dpH}$$

Equation 7.9.14

KSU IASRI CENTRAL UNIVERSITY LIBRARY

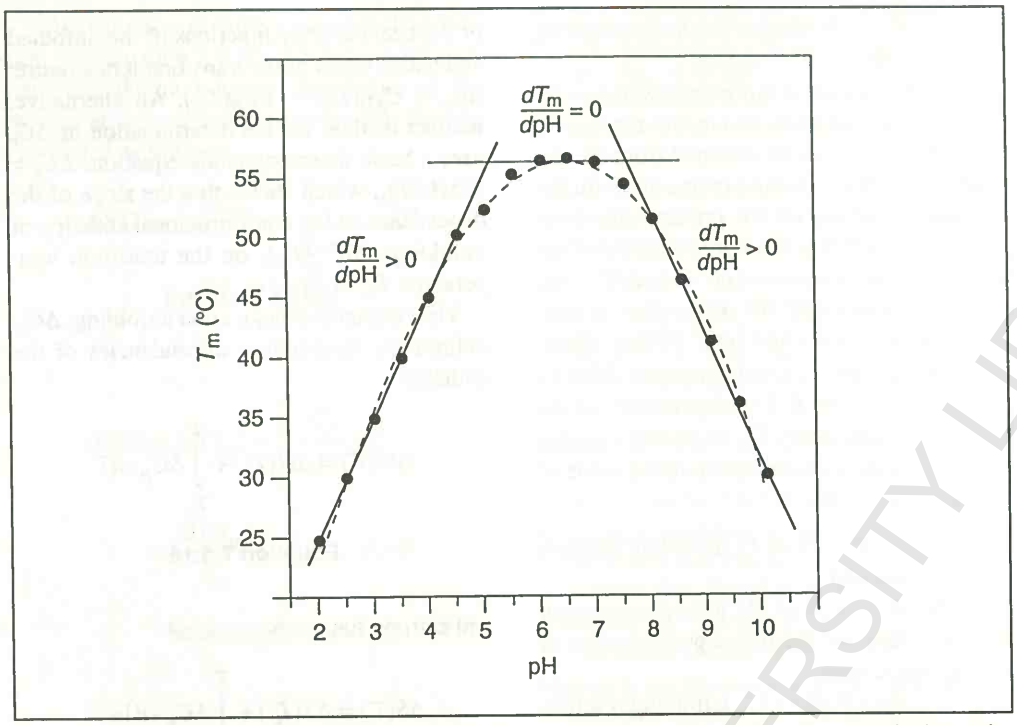


Figure 7.9.2 Dependence of the transition temperature on solution pH for a hypothetical protein.

Typically, the curve representing the dependence of the transition temperature on the pH of the solution has a bell shape (Fig. 7.9.2). On the left shoulder, $dT_m/d\text{pH} > 0$ and thus $\Delta\nu < 0$—i.e., the protein releases protons upon unfolding. On the right shoulder of T_m dependence on pH, $dT_m/d\text{pH} < 0$ and thus the protein absorbs protons upon unfolding ($\Delta\nu > 0$). At the maximum of the T_m(pH) function, $dT_m/d\text{pH} = 0$ and thus the protonation states of the native and unfolded states of a protein are identical ($\Delta\nu = 0$).

Van't Hoff enthalpy, $\Delta H_{vH}(T_m)$, and modes of unfolding transition. The so-called van't Hoff enthalpy for a two-state unfolding process is determined by the temperature dependence of the equilibrium constant, K_{eq}:

$$\Delta H_{vH}(T_m) = -R \cdot \frac{d \ln K_{eq}}{d(1/T)} = R \cdot T^2 \cdot \frac{d \ln K_{eq}}{dT}$$

Equation 7.9.15

where K_{eq} is defined as

$$K_{eq}(T) = \frac{F_N(T)}{F_N(T) + F_U(T)}$$

Equation 7.9.16

From this definition, it follows that van't Hoff enthalpy can be estimated from Equation 7.9.8 or Equation 7.9.9. However, for the calorimetric curves ΔH_{vH} (T_m) can be estimated as:

$$\Delta H_{vH}(T_m) = \frac{4 \cdot R \cdot T_m^2 \cdot < C_p(T_m) >^{max}}{\Delta H_{exp}(T_m)}$$

Equation 7.9.17

where $<C_p(T_m)>^{max}$ is the maximum value of the excess heat capacity function.

A direct estimate of two enthalpies—experimental calorimetric enthalpy, ΔH_{exp} (T_m), and van't Hoff enthalpy, $\Delta H_{vH}(T_m)$—from a single DSC experiment provides a unique way to determine the modes of protein unfolding (Biltonen and Freire, 1978; Privalov and Potekhin, 1986; Kidokoro and Wada, 1987; Marky and Breslauer, 1987; Ghosaini et al., 1988). When the ratio $\beta = \Delta H_{vH}(T_m)/\Delta H_{cal}(T_m)$ equals 1 (in practice, ~0.95 to 1.05), this indicates that the observed transition is two-state, proceeding from the native to the unfolded state without a significant population of intermediates: i.e., the unfolding reaction can be represented by N ↔ U. Deviations from unity indicate that the transition is more complicated. There are two possible explanations of β deviating from unity for

BCU IASRI CENTRAL UNIVERSITY LIBRARY

the reversible unfolding of the protein showing a single-isotherm DSC profile.

1. *The reaction is not monomolecular.* Information on whether or not the reaction is monomolecular can be obtained from the dependence of the transition temperature for the reversible reaction on the protein concentration. The possibility of oligomerization of the protein in the native or/and unfolded states should be examined by performing experiments under exactly the same solvent conditions using protein concentrations that differ by at least a factor of 5. The dependence of the transition temperature, T_m, on protein concentration, w, will be an indication of the mode of unfolding reaction:

a. $\partial T_m/\partial w = 0$: N $\leftrightarrow$ U, unfolding reaction is monomolecular;

b. $\partial T_m/\partial w > 0$: N $\leftrightarrow n$U, unfolding reaction is of nth order and the native protein exists in an oligomeric state;

c. $\partial T_m/\partial w < 0$: N $\leftrightarrow$ U_n, unfolding reaction is of nth order and the unfolded protein exists in an oligomeric state.

Situations (a) and (b) are very common and the thermodynamic formalism for their analysis is well developed. Situation (c) is rare for reversible reactions, but does occur for irreversible reactions when irreversibility is caused by aggregation (usually nonspecific) in the unfolded state.

2. *Intermediate states.* An alternative explanation for β being >1 is the existence of intermediate states during equilibrium reversible unfolding: i.e., N $\leftrightarrow$ $I_1 \leftrightarrow I_2$, ... $I_n \leftrightarrow$ U.

Consider the most common example of such an unfolding reaction, which is the unfolding of a two-domain protein. Both domains have similar intrinsic stability, $\Delta G_1(T) \approx \Delta G_2(T)$, and their interactions with one another are described by Gibbs free energy $\Delta G_{12}(T)$. The ratio of calorimetric to van't Hoff enthalpy, β, will depend on the strength of the interdomain interactions. If $\Delta G_{12}(T) = 0$, i.e., there is no interaction between domains, each domain will unfold independently of the other, and $\beta = 2$. If there is a strong interaction between domains, i.e., neither domain can exist in a folded state if the other domain is unfolded, unfolding will be very cooperative and $\beta = 1$.

Heat capacity change upon unfolding, ΔC_p, and thermodynamic functions for protein unfolding: enthalpy, $\Delta H(T)$, entropy, $\Delta S(T)$, and Gibbs free energy, $\Delta G(T)$. One way of estimating the heat capacity change upon protein unfolding is directly from the difference of the heat capacity functions of the unfolded and native states at the transition temperature: $\Delta C_p = C_{p,U}(T_m) - C_{p,N}(T_m)$. An alternative, indirect method for the determination of ΔC_p uses a basic thermodynamic equation, $\Delta C_p = (\partial \Delta H/\partial T)_p$, which shows that the slope of the dependence of the conformational enthalpy of unfolding, $\Delta H^{cnf}(T_m)$, on the transition temperature, T_m, is equal to ΔC_p.

Heat capacity change upon unfolding, ΔC_p, defines the temperature dependencies of the enthalpy

$$\Delta H(T) = \Delta H(T_0) + \int_{T_0}^{T} \Delta C_p \cdot dT$$

Equation 7.9.18

and entropy functions,

$$\Delta S(T) = \Delta S(T_0) + \int_{T_0}^{T} \Delta C_p \cdot d\ln T$$

Equation 7.9.19

and thus the Gibbs free energy function:

$$\Delta G(T) = -R \cdot T \cdot \ln K_{eq} = \Delta H(T) - T \cdot \Delta S(T)$$

Equation 7.9.20

The integration constants $\Delta H(T_o)$ and $\Delta H(T_o)$ in Equation 7.9.19 and Equation 7.9.20 have to be determined experimentally. The integration constant the enthalpy function can be set as $\Delta H(T_o) = \Delta H_{exp}(T_m)$. The integration constant for the entropy function (Equation 7.9.19) for a two-state transition is determined as follows. If the transition is two-state—i.e., $\Delta H_{exp}(T_m) \approx \Delta H_{vH}(T_m)$—then at the transition temperature (T_m) the population of native and unfolded states is equal. This means that at T_m the equilibrium constant of the reaction $K_{eq} = F_N/F_U$ is equal to 1, and thus the Gibbs free energy difference between the native and unfolded states at T_m is equal to zero and $\Delta G(T_m) = \Delta H(T_m) - T_m \cdot \Delta S(T_m) = 0$. Thus the entropy change upon protein unfolding at the transition temperature, $\Delta S(T_m)$, is equal to:

$$\Delta S(T_m) = \frac{\Delta H(T_m)}{T_m}$$

Equation 7.9.21

PCU IASS (CENTRAL UNIVERSITY LIBRARY)

This equation provides a constant of integration for Equation 7.9.20, $\Delta S(T_m) = \Delta S(T_o)$. A complete thermodynamic description of the system at *any* other temperature can then be calculated as follows:

$$\Delta H(T) = \Delta H(T_m) + \int_{T_m}^{T} \Delta C_p \cdot dT$$

Equation 7.9.22

$$\Delta S(T) = \Delta S(T_m) + \int_{T_m}^{T} \frac{\Delta C_p}{T} \cdot dT$$

$$= \frac{\Delta H(T_m)}{T_m} + \int_{T_m}^{T} \Delta C_p \cdot d\ln T$$

Equation 7.9.23

$$\Delta G(T) = \Delta H(T_m) - T \cdot \frac{\Delta H(T_m)}{T_m}$$

$$+ \int_{T_m}^{T} \Delta C_p \cdot dT - T \cdot \int_{T_m}^{T} \Delta C_p \cdot d\ln T$$

Equation 7.9.24

where the heat capacity change upon protein unfolding, ΔC_p, is a function of temperature itself. In practice, however, if no extrapolations over a wide temperature range are necessary, as a first approximation a temperature-independent ΔC_p is a reasonable assumption.

Time Considerations

Sample preparation (of which dialysis is the most time-consuming step) usually requires 12 to 16 hr. Consult *UNIT 4.4* for possible ways of reducing dialysis time. Baseline data is usually collected in overnight experiments; a calorimetric experiment with a protein sample requires several hours, depending on the heating rate. However, optimization of experimental conditions (reversibility and equilibrium) when first working with an unfamiliar sample may take several days. Time requirements for data analysis vary in individual cases.

Literature Cited

Biltonen, R.L. and Freire, E. 1978. Thermodynamic characterization of conformational states of biological macromolecules using differential scanning calorimetry. *CRC Crit. Rev. Biochem.* 5:85-124.

Ghosaini, L.R., Brown, A.M., and Sturtevant, J.M. 1988. Scanning calorimetric study of the thermal unfolding of catabolite activator protein from *Escherichia coli* in the absence and presence of cyclic mononucleotides. *Biochemistry* 27:5257-5261.

Gill, S.C. and von Hippel, P.H. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182:319-326.

Izatt, R.M. and Christensen, J.J. 1976. Heats of proton ionization, pK, and related thermodynamic quantities. *In* CRC Handbook of Biochemistry and Molecular Biology, Vol. 1: Physical and Chemical Data (G.D. Fasman, ed.) pp. 151-269. CRC Press, Cleveland.

Jaenicke, L. 1974. A rapid micromethod for the determination of nitrogen and phosphate in biological material. *Anal. Biochem.* 61:623-627.

Kidokoro, S. and Wada, A. 1987. Determination of thermodynamic functions from scanning calorimetry data. *Biopolymers* 26:213-229.

Krishinan, K.S. and Brandts, J.F. 1978. Scanning calorimetry. *Methods Enzymol.* 49:3-14.

Mabrey, S. and Sturtevant, J.M. 1976. Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proc. Natl. Acad. Sci. U.S.A.* 73:3862-3866.

Makhatadze, G.I. 1998. Heat capacities of amino acids, peptides and proteins. *Biophys. Chem.* 71:1-26.

Makhatadze, G.I., Medvedkin, V.N., and Privalov, P.L. 1990. Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range. *Biopolymers* 30:1001-1010.

Marky, L.A. and Breslauer, K.J. 1987. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* 26:1601-1620.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4:2411-2423.

Plotnikov, V.V., Brandts, J.M., Lin, L.N., and Brandts, J.F. 1997. A new ultrasensitive scanning calorimeter. *Anal. Biochem.* 250:237-244.

Privalov, G., Kavina, V., Freire, E., and Privalov, P.L. 1995. Precise scanning calorimeter for studying thermal properties of biological macromolecules in dilute solution. *Anal. Biochem.* 232:79-85.

Privalov, P.L. and Potekhin, S.A. 1986. Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.* 131:4-51.

Privalov, P.L. and Khechinashvili, N.N. 1974. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.* 86:665-684.

Sanchez-Ruiz, J.M. 1992. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys. J.* 61:921-935.

BOULDER CENTRAL UNIVERSITY LIBRARY

Scopes, R.K. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* 59:277-282.

Sturtevant, J.M. 1972. Calorimetry. *Methods Enzymol.* 26:227-253.

Winder, A.F. and Gent, W.L. 1971. Correction of light-scattering errors in spectrophotometric protein determinations. *Biopolymers* 10:1243-1251.

Wintrode, P.L., Makhatadze, G.I., and Privalov, P.L. 1994. Thermodynamics of ubiquitin unfolding. *Proteins* 18:246-253.

Yu, Y., Makhatadze, G.I., Pace, C.N., and Privalov, P.L. 1994. Energetics of ribonuclease T1 structure. *Biochemistry* 33:3312-3319.

Key References

Biltonen and Freire, 1978. See above.

Privalov and Potekhin, 1986. See above.

Kidokoro and Wada, 1987. See above.

Three references that discuss in detail the theoretical background for deconvolution of the heat capacity profile, each describing a separate approach to data analysis.

Izatt and Christensen, 1976. See above.

Basic reference containing tables of heats of ionizations and other thermodynamic information for different buffer systems.

Pace et al., 1995. See above.

Detailed discussion on the calculation of protein extinction coefficients from amino acid compositions.

Privalov, P.L. 1979. Stability of proteins: Small globular proteins. *Adv. Protein Chem.* 33:167-241.

Privalov, P.L. 1982. Stability of proteins: Proteins which do not present a single cooperative system. *Adv. Protein. Chem.* 35:1-104.

Makhatadze, G.I. and Privalov, P.L. 1995. Energetics of protein structure. *Adv. Protein. Chem.* 47:307-425.

Three reviews describing and assessing the determination and interpretation of protein stability by DSC.

Sanchez-Ruiz, 1992. See above.

Detailed discussion of the analysis of irreversible denaturation processes.

Contributed by George I. Makhatadze
Texas Tech University
Lubbock, Texas

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CONVENTIONAL CHROMATOGRAPHIC SEPARATIONS

8

INTRODUCTION 8.0.1

8.1 Overview of Conventional Chromatography **8.1.1**
Yield and Purity of Target Protein 8.1.1
Steps in a Purification Strategy 8.1.3
Parameters Affecting Chromatographic Resolution 8.1.5
Literature Cited 8.1.9

8.2 Ion-Exchange Chromatography **8.2.1**
Strategic Planning 8.2.2
Basic Protocol 1: Batch Adsorption and Step-Gradient Elution with Increasing Salt Concentration 8.2.10
Alternate Protocol: pH-Based Step-Gradient Elution 8.2.13
Basic Protocol 2: Column Chromatography with Linear Gradient Elution 8.2.14
Support Protocol 1: Test Tube Pilot Experiment to Determine Starting Conditions for Ion-Exchange Chromatography 8.2.16
Support Protocol 2: Measurement of Dyamic (Column) Capacity and Breakthrough Capacity of Ion-Exchange Columns 8.2.20
Support Protocol 3: Gradient-Formation Techniques 8.2.21
Support Protocol 4: Cleaning and Regeneration of Ion-Exchange Media 8.2.23
Support Protocol 5: Storage of Ion-Exchange Media 8.2.24
Commentary 8.2.25

8.3 Gel-Filtration Chromatography **8.3.1**
Strategic Planning 8.3.1
Basic Protocol 1: Desalting (Group Separation) 8.3.6
Basic Protocol 2: Protein Fractionation 8.3.14
Basic Protocol 3: Determination of Molecular Size 8.3.17
Support Protocol: Column Calibration 8.3.19
Reagents and Solutions 8.3.22
Commentary 8.3.24

8.4 Hydrophobic-Interaction Chromatography **8.4.1**
Strategic Planning 8.4.2
Basic Protocol 1: Packing a Column with an HIC Matrix of ≥90-mm Beads 8.4.6
Alternate Protocol: Packing a Column with an HIC Matrix of ≤34-mm Beads 8.4.8
Basic Protocol 2: Testing the Packed Bed 8.4.8
Support Protocol 1: Pilot Experiment to Choose a Matrix and Determine Binding and Elution Conditions for HIC 8.4.10
Basic Protocol 3: Elution of Proteins from HIC Columns 8.4.11
Support Protocol 2: Regeneration, Cleaning, and Storage of HIC Columns 8.4.12
Commentary 8.4.15

8.5 Chromatofocusing **8.5.1**
Strategic Planning 8.5.1
Basic Protocol: Separation of Proteins by Chromatofocusing 8.5.3
Support Protocol 1: Regeneration of the CF Column 8.5.7
Support Protocol 2: Removal of Polybuffer from Protein 8.5.8
Commentary 8.5.8

continued

BCU IAS CENTRAL UNIVERSITY LIBRARY

Copyright © 1997 by John Wiley & Sons, Inc.

8.6 Hydroxylapatite Chromatography **8.6.1**
Strategic Planning 8.6.1
Basic Protocol 1: Low-Pressure Chromatography for Protein Purification 8.6.1
Alternate Protocol 1: Simple Column Chromatography for Protein Purification Using Hydroxylapatite 8.6.6
Alternate Protocol 2: Batch Processing for Protein Purification Using Hydroxylapatite 8.6.6
Basic Protocol 2: HPLC of Proteins and Nucleic Acids Using Hydroxylapatite 8.6.7
Commentary 8.6.9

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 8
Conventional Chromatographic Separations

INTRODUCTION

It can be argued with conviction that the purification of proteins has been a key step in many of the advances in our understanding of biological structure and function. The careful dissection of enzyme mechanisms would have been impossible without isolated catalytic species. Before the advent of gene sequencing, determination of protein structure absolutely required clean protein samples. Precise determination of the biological effects of cytokines and growth factors necessitates the preparation of pure polypeptides. Impressive advances from X-ray structure determination must begin with homogeneous, concentrated protein samples. The utility of site-specific mutagenesis, which is crucial to asking precise questions about structure-function relationships, is limited if the recombinant protein cannot be separated from contaminating cellular proteins. Finally, the revolution created by the availability of restriction enzymes is founded upon the preparation of pure samples of the reagents.

This chapter presents a rational, systematic approach to the discovery and optimization of protein separations by chromatographic methods. The first unit provides a description of how to plan an overall strategy to attempt the isolation of a new protein (*UNIT 8.1*). Important initial considerations include the desired quality level of the final product and the properties of the source as well as the anticipated properties of the target protein. The choice of a strategic approach is described, beginning with consideration of the needs of the final sample with respect to concentration, buffer composition, and pH. The competing factors of selectivity, efficiency, and capacity are discussed with reference to how each method can influence the choice of the next step.

The second unit provides a series of steps to identify conditions for optimizing a separation strategy (*UNIT 8.2*). Alternative methods for sample processing and elution are described, thus providing numerous choices for the fractionation of a variety of proteins. In addition, critical parameters for the long-term care of a chromatography system are detailed. The relationship of the ion-exchange method to other procedures is also discussed.

A single procedure, no matter how carefully performed, is unlikely to totally resolve a mixture of proteins. Accordingly, use of additional procedures that provide separation based on different physical properties are essential to complete the isolation procedure. Protein fractionation by gel-filtration chromatography can provide a separation based on molecular size (*UNIT 8.3*). Gel filtration can also be used to remove salts from a protein at various stages of isolation; this may be valuable before ion-exchange chromatography, for example. When performed using standard proteins for calibration, gel filtration can also provide information about the size of a new protein.

Another orthogonal separation method that has proven to be very valuable, particularly for protein isolation, is hydrophobic interaction chromatography (*UNIT 8.4*). In this method the protein interacts with a matrix through surface-exposed hydrophobic groups. Guidelines for beginning and optimizing a separation based on hydrophobic interaction chromatography are provided.

BOSTON CENTRAL UNIVERSITY LIBRARY

Contributed by Ben M. Dunn
Current Protocols in Protein Science (1997) 8.0.1-8.0.2
Copyright © 1997 by John Wiley & Sons, Inc.

Chromatofocusing (*UNIT 8.5*) is a technique that separates proteins on the basis of their isoelectric points. High resolution is achieved, making the procedure perfect for intermediate polishing steps. Selection of proper media and buffers is essential to the technique, as described in the unit.

Binding of the charged groups of proteins (positive as well as negative) to the calcium phosphate matrix of hydroxylapatite provides another unique separation mechanism to resolve protein mixtures (*UNIT 8.6*). The unit also includes recommendations for alternative methods to derive the optimal advantages of this strategy.

Each unit defines chromatographic parameters and describes methodology to provide quantitative evaluation of column efficiency. In addition, each unit presents summaries of the properties of chromatographic media available from commercial sources, which have undergone significant improvements over the past twenty years. The advantages of some features such as pore size and bead diameter are discussed in the context of the competing needs of resolution, capacity, and speed.

Although all the procedures are presented so that separations can be done manually, the advantages of automation for some steps are described. As separation technology is rapidly changing, future updates will expand the methodology described in this chapter.

Ben M. Dunn

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Overview of Conventional Chromatography

The key to success in purification of any protein is developing the right separation strategy. This involves selecting and correctly applying a combination of separation techniques, chosen on the basis of information about the target protein and contaminants. There is no single separation technique and no single purification scheme that will allow successful purification of all types of proteins. However, there is a single approach to developing a successful purification strategy.

Three basic steps must be completed before beginning to develop a purification strategy. First, the amount and purity of the target protein (the protein to be purified) required for its intended use must be determined. Second, all available physical and chemical information concerning potential source materials, the nature of the target protein itself, and major contaminants must be gathered. Third, the best source material from which to purify the target protein must be determined. These three steps are mutually interdependent—what is learned in one will have an impact on the others. The more that is known before beginning development of a purification strategy, the more successful that strategy will be.

YIELD AND PURITY OF TARGET PROTEIN

How much protein must be purified? How pure must the protein be? These two questions establish the criteria for selecting specific chromatographic techniques and for judging the efficacy of any individual purification step or overall purification procedure. The answers to these simple questions depend upon the nature of the source material from which the target protein will be purified, the intended use for the purified protein, and the methods available to assess the purity of the target protein.

Amount of Protein Required

The amount of protein required from a purification process depends primarily on the intended use for the target protein. However, the amount of protein that will be used during the purification process to develop methods and assess purity must also be considered. It is necessary only to estimate whether analytical (≤1 mg), preparative (≤1 g), pilot (≤ 1 Kg), or process (>1 Kg) amounts of target protein will be needed. This is often referred to as the scale of operation. When the source material is very expensive, unstable, or difficult to obtain, it is necessary to develop and optimize the purification strategy on a small scale first, then scale up. Conversely, if multiple samples are to be processed, as for analytical applications, it may be desirable to develop the purification process at a larger scale, then scale down. The amount of target protein to be purified may also be determined by limited availability of source material.

Source Material

Once the scale of operation has been determined, the amount of source material needed to begin the procedure must be clearly defined. This obviously depends on the concentration of target protein in the source material, which can vary greatly. Certain proteins and peptides are present in natural sources only in very minute quantities. Purification of even a microgram of pure material may require many kilograms of tissue, or many liters of fermentation broth if microorganisms are the source. On the other hand, some recombinant proteins are expressed at such high levels in a host that the protein may be present in milligram per milliliter or higher concentrations. Source materials containing very low or very high concentrations of target protein each present unique problems that must be addressed in selecting a purification strategy.

Yield of Purification Procedure

The next major consideration in deciding how much protein to purify is what yield is necessary at each step of the purification procedure to ensure that there will be enough pure target protein at the end, taking into account the availability of source material. If a 75% yield of target protein is obtained at each of four steps in a purification procedure, less than one-third of the desired protein will remain at the end of the procedure (Fig. 8.1.1). It is obviously advantageous to use as few steps as possible in a purification process while maintaining as high a yield as possible at each step.

The yield from any purification step may be reported as total protein, percentage of total activity, or change in specific activity. The yield information from progressive purification steps is used to construct a purification table (e.g., see Table 8.1.1).

In Table 8.1.1, the yield is reported as percent of total activity remaining from the crude

Copyright © 1995 by John Wiley & Sons, Inc.

RUTGERS / CAMDEN UNIVERSITY LIBRARY

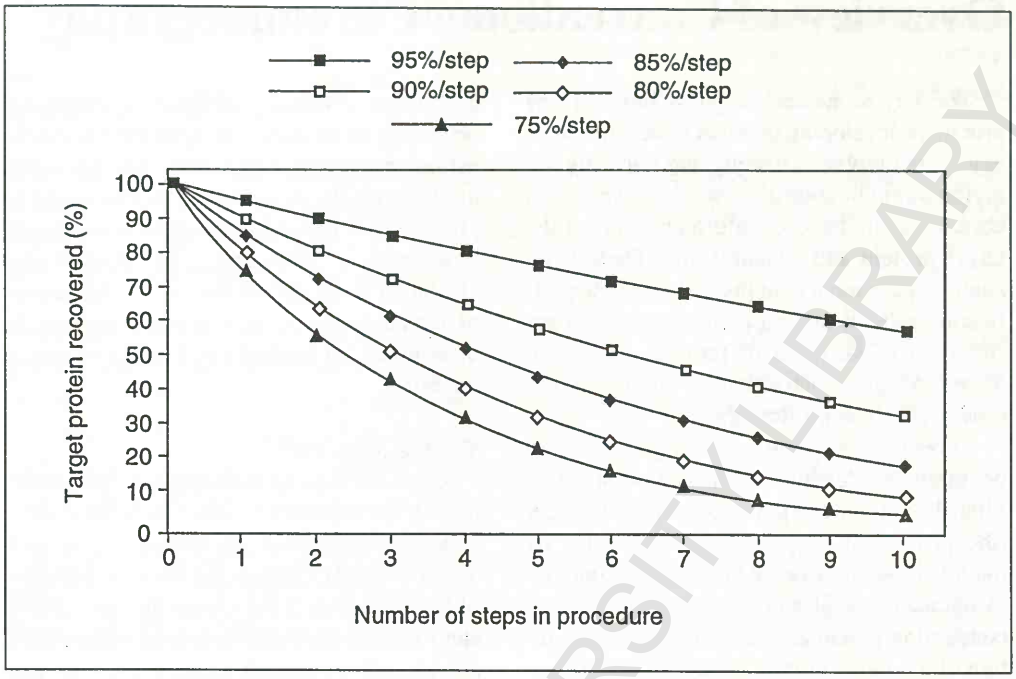


Figure 8.1.1 Theoretical yields from multistep protein purifications. Each curve represents a multistep process with a given percentage yield per step. The percentage of the original quantity of target protein remaining at the end of purification is plotted against the number of purification steps.

homogenate; the step yield is the percent activity relative to the previous step; and the purification factor reflects the change in specific activity relative to the crude homogenate. It is not unusual to have yields >100% based on activity (e.g., the yield after anion-exchange chromatography in Table 8.1.1). This is most often due to the removal of inhibitors or contaminants that interfere with the activity assay.

Assessment of Protein Purity

Assessing protein purity is often the most time-consuming and labor-intensive step in the purification process. It is generally desirable to limit the number of steps in a purification process simply to reduce the number of assays that must be performed to assess purity at each step. Protein purity is judged from a structural and/or functional perspective, depending on the nature of the target protein and its intended use. If the intention is to examine the function of the target molecule, the activity of the target protein must be maintained during the purification process. A functional assessment of the purity of the target protein will thus be required at each step. If the intention is to examine the structure of the target molecule, structure must be maintained, but function may or may not have to be preserved. A structural assessment of purity will, however, be required in this case. If structural/functional relationships involving the target molecule are to be examined, assessment of both structural and functional purity will be required.

Methods for assessing protein purity must quantitate the amount of target protein relative to the amount of contaminant(s) in the sample. This requires two separate analytical methods: one for the target protein and one for the total protein including contaminants. For example, with enzymatic proteins, assessment of functional purity typically relies on kinetic assays in which the amount of substrate consumed or product produced per unit time is proportional to the amount of enzyme present. The total activity of the sample is calculated and compared to the total amount of protein to give the specific activity of the sample. Methods for determining total protein are discussed in Chapter 3. An increase in specific activity at a purification step reflects a loss of contaminant proteins.

Functional purity of structural proteins, denatured proteins (i.e., proteins that have lost their function during the purification procedure), and proteins for which a functional assay does not exist must usually be assessed by measuring structural attribute(s)—e.g., molecular weight, pI, presence of a metal ion or cofactor, or presence of an antibody binding

PCU IASRI CENTRAL UNIVERSITY LIBRARY

Table 8.1.1 Purification Table for an Arbitrary Enzyme[a]

Purification step	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Yield (% total activity)	Step yield[b]	Purification factor[c]
Crude homogenate	200	50,000	0.004	100	—	1.00
12,000 × *g* supernatant	180	35,000	0.005	90	90	1.25
$(NH_4)_2SO_4$ fraction (20-50%)	150	25,000	0.006	75	83	1.50
HIC	125	2500	0.05	62.5	83	12.50
AEX	208	29.3	7.1	104	166	1775
Gel filtration	148	5.9	25.1	74	71	6275

[a]Abbreviations: HIC, hydrophobic-interaction chromatography; AEX, anion-exchange chromatography.

[b]Percentage activity relative to previous step.

[c]Increase in specific activity.

site. As in the case of functional purity, an increase in the amount of target protein relative to total protein reflects an increase in purity.

STEPS IN A PURIFICATION STRATEGY

A purification strategy for any protein at any scale of operation can be broadly divided into three sequential stages—capture, intermediate purification, and polishing. Each stage represents a set of specific problems that may be encountered during a purification process. The nature of the sample and the scale of operation will dictate what equipment and methodology are appropriate to solve the problem.

Capture Stage

The capture stage is the initial purification of the target protein from the source material. The goal of capture is to concentrate the target protein while removing as much of the major contaminant(s) as possible; to this end, an adsorptive chromatographic technique should be employed. Ion-exchange chromatography (*UNIT 8.2*) and hydrophobic-interaction chromatography (HIC; *UNIT 8.4*) are generally the best chromatographic techniques for capture at any scale of operation. Certain affinity techniques are also effective, including lectin affinity chromatography (*UNIT 9.1*), dye affinity chromatography (*UNIT 9.2*), immunoaffinity chromatography, and metal-chelate affinity chromatography (MCAC). Affinity chromatography is very useful for capture on a small scale (e.g., <1 g). Larger-scale capture using affinity-based methods is generally avoided because of the cost of the chromatographic media and reagents (especially for immunoaffinity chromatography).

At the capture stage the sample is very crude. It may be turbid from the presence of cellular debris or viscous from the presence of DNA, and will typically represent the largest volume to be processed during the purification procedure. Water is often the major contaminant, especially in samples from cell-culture fluids or fermentation broth. The sample may also contain pigments, particulates, and lipophilic materials that can easily clog chromatography columns or bind nonspecifically to certain media. Once such contaminants are introduced into a chromatography system, they can be very difficult to remove. The sample may also contain proteinases or other agents that can potentially destroy or denature the target protein.

Ion-exchange or HIC media suitable for capture must allow high throughput of dilute, turbid, or viscous samples. This is achieved by using bead diameters of 40 to 300 µm. Columns packed with such relatively large beads can usually be operated at lower pressures and higher flow rates, and are less prone to fouling or clogging from particulates in the sample. Higher flow rates reduce the time required to handle large volumes, which in turn reduces the potential for product loss from degradation or denaturation.

When the volume of crude sample is very large (or larger than can be reasonably handled with available pumps or columns), batch techniques (*UNIT 8.2*) can be used at the capture stage. In batch adsorption of the target protein by ion exchange, HIC, or affinity chromatography, the chromatography medium is mixed directly with the sample in a suitable vessel. When the target protein is adsorbed, the medium is separated from the sample by filtration, centrifu-

POLK CENTRAL UNIVERSITY LIBRARY

gation, or settling and decantation. The target protein is then eluted from the medium and collected. For both column and batch techniques, elution of the target protein at the capture stage is often achieved using simple step gradients, but continuous gradients may be used to achieve higher resolution with column techniques. Specific medium requirements and methodologies for use at the capture stage are presented in *UNIT 8.2* for ion-exchange chromatography and in *UNIT 8.4* for HIC.

Intermediate Purification Stage

The next stage in the purification process is intermediate purification. Once the solution containing the target protein has been clarified and concentrated, only small amounts of other types of biomolecules (e.g., lipids and nucleic acids) should be present. The target protein and remaining protein contaminants have a functional or structural attribute in common, whose nature depends on the technique that was used at the capture stage. If ion exchange was used, the contaminants and target protein will be similar in charge characteristics; if HIC was used, similar in hydrophobicity; if affinity methods were used, similar in a ligand- or antibody-binding site, and if ultrafiltration or gel filtration was used, similar in size.

Because the contaminants have a physicochemical similarity to the target protein, it is best to exploit a different attribute of the target protein or contaminant during the intermediate purification stage than was used during the capture stage. Use of the same chromatographic technique for both capture and intermediate purification is not recommended. The combination of ion exchange and HIC is very effective in most purification strategies. Proteins are eluted from HIC media at low ionic strength, which is ideal for direct application of the sample to an ion-exchange column, where binding occurs at low ionic strength. Conversely, proteins are eluted from ion-exchange media with increasing ionic strength, which is ideal for subsequent direct application of the sample to a HIC column, where binding occurs at high ionic strength. Cation and anion exchange in combination can also be very effective. The reduced sample volume in the intermediate purification stage also allows use of negative adsorption chromatography, which is designed to bind contaminants to the column, allowing the target protein to pass through.

Chromatography techniques used in intermediate purification should resolve the target protein from contaminants on the basis of small differences in a single physicochemical attribute. High resolution and high yield are the primary goals of intermediate purification. To achieve these goals, a chromatography medium of smaller bead size than that used at the capture stage should be selected. The advantage of smaller bead diameters is increased resolution resulting from decreased peak width. However, smaller bead diameters typically require higher operating pressures. To further enhance resolution during intermediate purification, elution of the target protein is often done using a linear gradient. Specific medium requirements and methodologies for use at the intermediate purification stage are presented in *UNIT 8.2* for ion-exchange chromatography and in *UNIT 8.4* for HIC.

Polishing Stage

Polishing is the final stage in the purification process. The goal of polishing is to remove structural and functional variants of the target protein, as well as any trace contaminants. After polishing, the target protein should be in a suitable form for its intended use. The contaminants and structural variants that must be removed during polishing share very similar physicochemical attributes with the target protein. Chromatographic techniques employed in the polishing stage should exploit a different physicochemical attribute of the target protein than those used for capture and intermediate purification. The similarity of contaminants to target protein heightens the need for high-resolution techniques during this stage. Chromatofocusing, gel filtration (*UNIT 8.3*), reversed-phase chromatography (RPC; Corran, 1989), and affinity-based techniques (Chapter 9) are all useful in this stage. RPC is very effective for resolving structural variants of a target protein, but generally leads to loss of activity of the target protein because of the organic solvents employed. Adsorptive techniques such as chromatofocusing and affinity chromatography often require special eluants to recover the target protein, which may have to be removed before the purified target protein can be used. Gel filtration does not generally require any special eluants, and is often the best technique for polishing.

Optimizing Purification

Throughout the three stages of the purification process, the concentration of target protein is increasing with respect to the concentration of contaminants. In addition, the contaminants to be removed at each stage are increasingly

BCU IASE - BETTAH UNIVERSITY LIBRARY

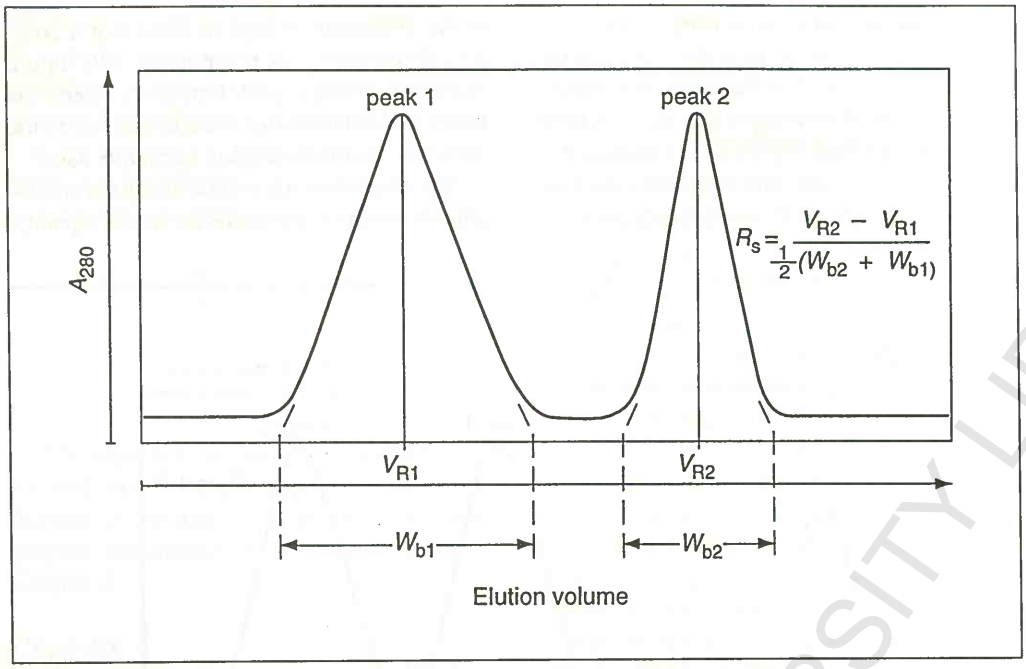


Figure 8.1.2 Hypothetical chromatogram illustrating calculation of elution volumes and peak widths for two peaks. Total elution volume is plotted against absorbance at 280 nm. V_{R1} and V_{R2} are the elution volumes (at the peak maxima) of peaks 1 and 2, respectively; W_{b1} and W_{b2} are the peak widths (at base) of peaks 1 and 2, respectively. Peak widths at base represent the segment of the baseline intercepted by the tangents drawn to the inflection points of the peak. Resolution (R_s) is defined as the distance between the elution volumes at peak maxima divided by the average of the peak widths, according to Equation 8.1.1 (see discussion of Resolution).

similar to the target protein. This increases the need for higher resolution and better yields at each progressive step. It is desirable to examine the efficacy of several techniques at each stage of the procedure before deciding on a specific one to use. Once a specific chromatographic technique has been chosen, the methodology for that step should be optimized: first to achieve maximum resolution, then maximum capacity and yield, and finally maximum speed. This approach will result in a fast, efficient, economical, and robust purification strategy.

PARAMETERS AFFECTING CHROMATOGRAPHIC RESOLUTION

Resolution

Resolution is a measure of the relative separation achieved between two chromatographically distinct materials, and maximum resolution is the primary goal of any purification step. This discussion covers the main theoretical parameters that affect ability to resolve components of a protein sample into chromatographically distinct forms, thereby providing a basis on which to judge chromatographic results. The main parameters affecting resolution are selectivity, efficiency, and capacity. More detailed discussions are found in Giddings and Keller (1965) and Janson and Ryden (1989).

The result of any chromatographic separation is often expressed as the resolution between the zones or peaks containing the chromatographically distinct sample components. The resolution (R_s) of two peaks obtained in a chromatography step may be determined from the chromatogram shown in Figure 8.1.2 according to Equation 8.1.1, where V_{R2} is the elution volume of peak 2, V_{R1} is the elution volume of peak 1, W_{b1} is the width of peak 1, and W_{b2} is the width of peak 2. In cases where a constant eluant flow is maintained and fractions are collected at regular time intervals, V_{R1}, V_{R2}, W_{b2}, and W_{b1} may be expressed in units of time rather than volume. It is necessary that the same units (time or volume) be used to express all terms in the equation to give a dimensionless value to R_s.

$$R_s = \frac{V_{R2} - V_{R1}}{\frac{1}{2}(W_{b2} + W_{b1})}$$

Equation 8.1.1

POLITECNICA UNIVERSITY LIBRARY

Resolution values can be used to determine if further optimization of a chromatographic procedure is required, as they provide a numerical indication of whether two peaks have been completely resolved. Figure 8.1.3 compares two separation results: one with completely resolved peaks and the other with incompletely resolved peaks. It should be kept in mind that a completely resolved peak is not necessarily a pure substance. A single peak frequently represents multiple substances that are not resolved by the particular chromatographic technique used.

The resolution achievable in any chromatographic system is proportional to the capacity,

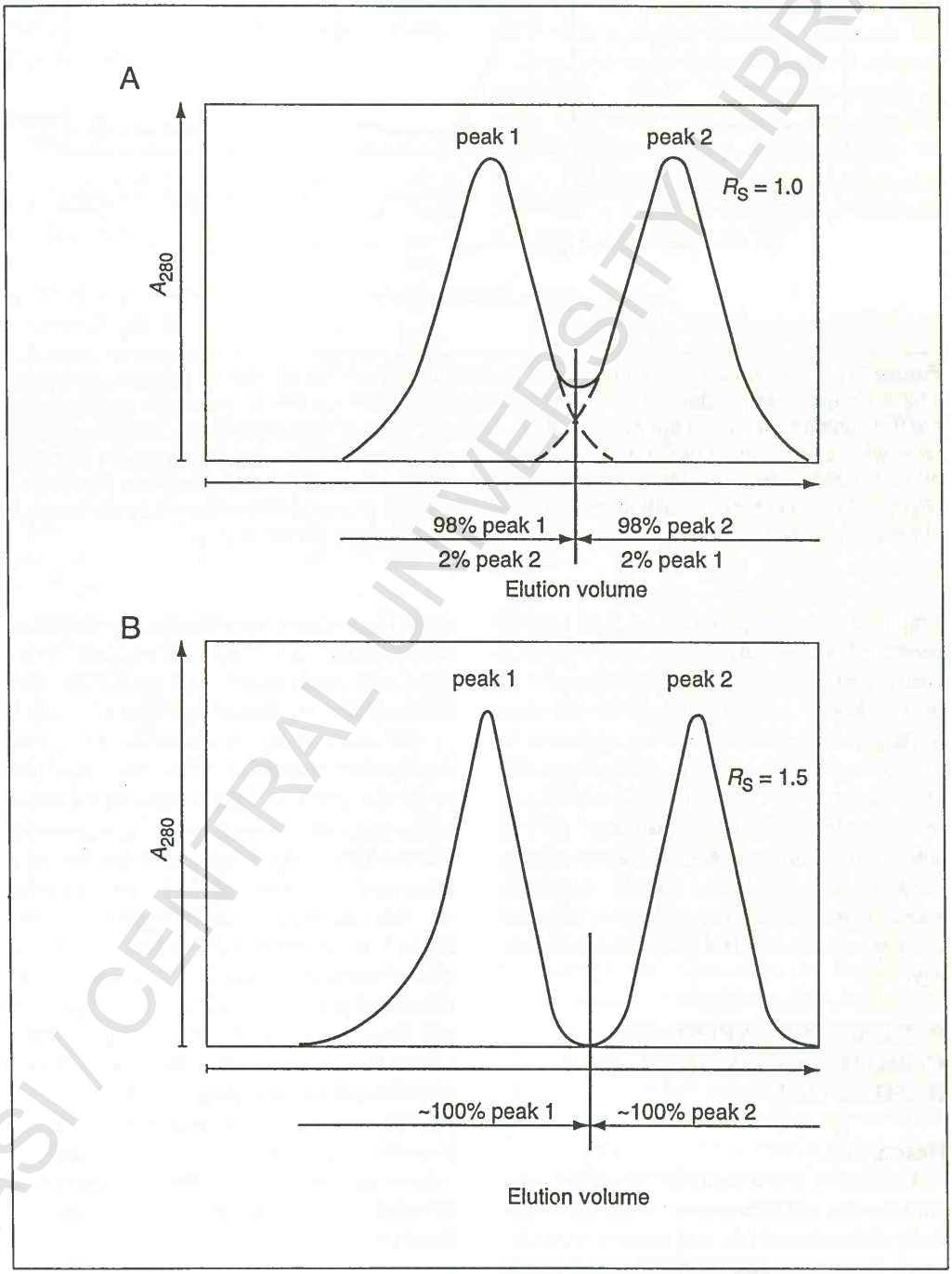


Figure 8.1.3 Separation results with different resolutions. **(A)** When R_s = 1.0, 98% purity has been achieved—i.e., 98% of the fraction collected at peak 1 consists of a single protein with the remaining 2% being contamination from the protein collected at peak 2; conversely, 98% of the fraction collected at peak 2 consists of a single protein with the remaining 2% being contamination from peak 1. **(B)** Complete (baseline) resolution requires that R_s >1.5. At this value, the peaks are completely separated from each other (i.e., the purity of each peak is >99.9%).

PCU IASI / CENTRAL UNIVERSITY LIBRARY

efficiency, and selectivity of the system. Each of these factors must be considered and controlled to achieve success. The theoretical expression for resolution is Equation 8.1.2, where k is the average capacity factor for the two peaks, N is the efficiency factor for the system, and α is the selectivity factor of the medium.

$$R_s = \frac{1}{4}\left(\frac{k}{1+k}\right)\left(N^{-2}\right)\left(\frac{\alpha-1}{\alpha}\right)$$

Equation 8.1.2

The impact of various experimental factors on resolution in various stages of protein purification is discussed for specific chromatographic techniques in subsequent units of Chapter 8.

Capacity

The capacity or retention factor (k) is a measure of retention of a sample component. It should not be confused with the loading capacity of a column, which is expressed as milligrams of sample bound per milliliter of gel and represented by the area under a peak. The capacity factor may be calculated for any individual peak in a chromatogram. For example, the capacity factor for peak 2 in Figure 8.1.4 is derived from Equation 8.1.3, where V_{R2} is the elution volume of peak 2 and V_m is the volume of the mobile phase (i.e., the total bed volume).

$$k = \frac{V_{R2} - V_m}{V_m}$$

Equation 8.1.3

In Equation 8.1.2 for R_s, presented in the discussion of Resolution, k is the average of k_1 (capacity factor for peak 1) and k_2 (capacity factor for peak 2).

Adsorption techniques such as ion exchange, HIC, chromatofocusing, RPC, and affinity chromatography can have high capacity factors because experimental conditions can be manipulated so that the elution volume for a peak can exceed the total bed volume (V_m as is the case with peaks 2 and 3 in Fig. 8.1.4.) However, in gel filtration, which is a non-adsorptive technique, all peaks must elute within the volume V_m–V_o as is the case with peak 1 in Fig. 8.1.4. The impact of experimental conditions on capacity at various stages of protein purification are discussed for different purification techniques in subsequent units of Chapter 8.

Efficiency

The efficiency factor (N) is a measure of zone broadening (peak width) occurring on a column. It can be calculated for any given peak from Equation 8.1.4, where V_R is the elution volume (i.e., the total volume of eluant that has

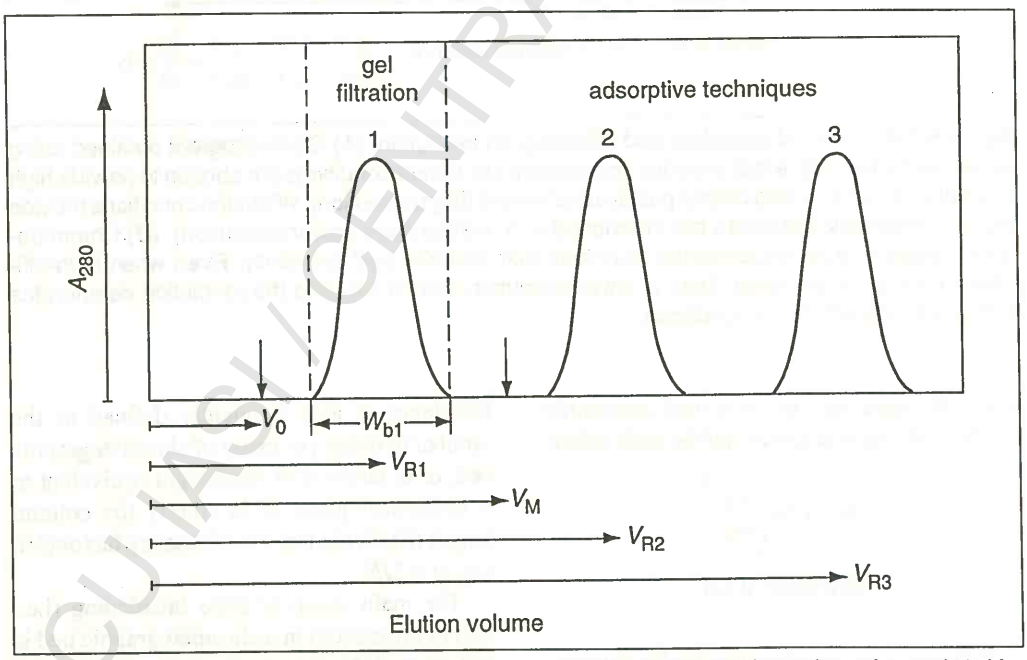


Figure 8.1.4 Hypothetical chromatogram. V_o = void volume; V_{R1} = elution volume for peak 1; V_{R2} = elution volume for peak 2; V_{R3} = elution volume for peak 3; V_M = volume of mobile phase; W_{b1} = peak width for peak 1; W_{b2} = peak width for peak 2.

FCU IASL / CENTRAL UNIVERSITY LIBRARY

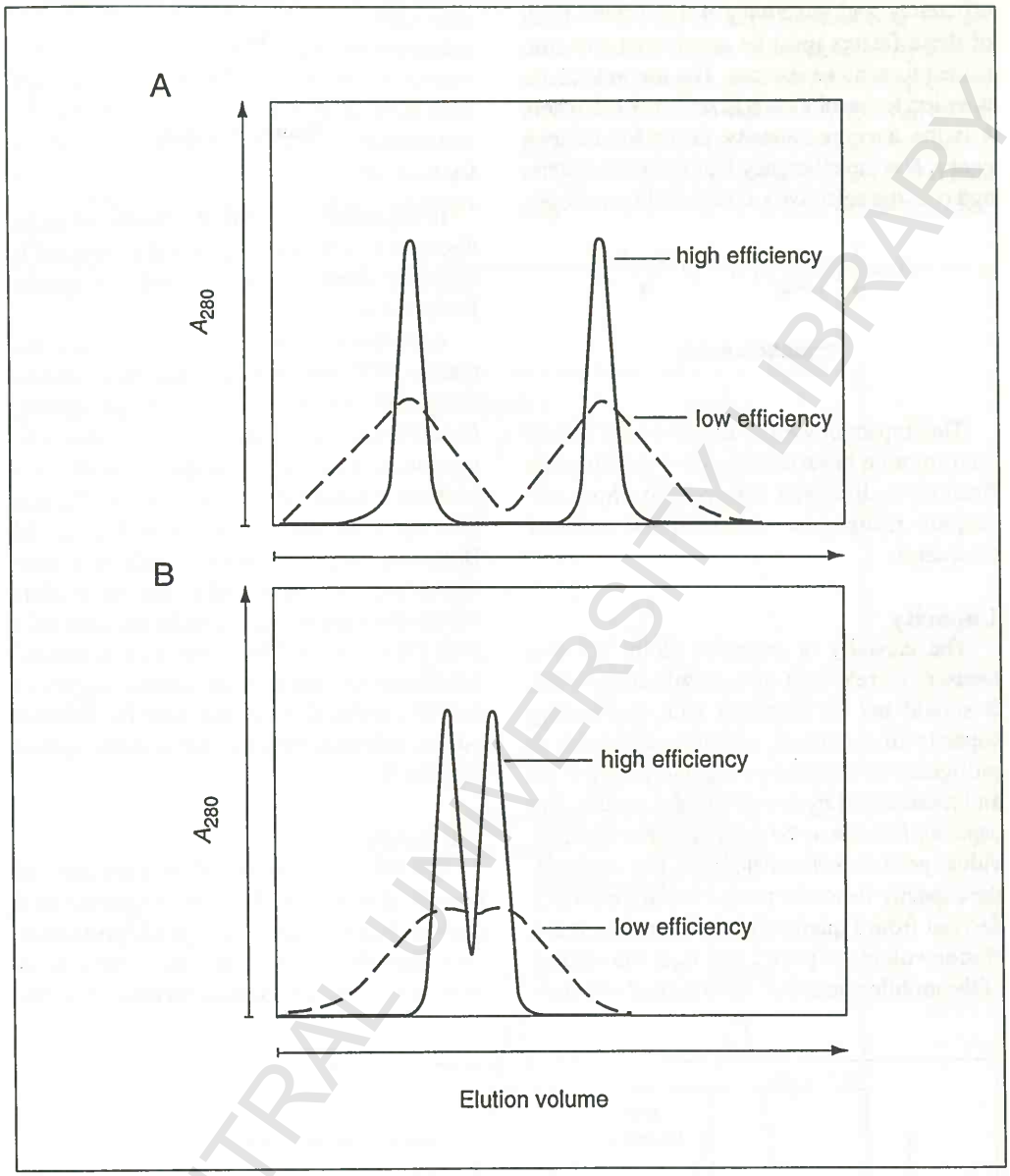


Figure 8.1.5 Effect of selectivity and efficiency on resolution. (**A**) Chromatogram obtained using experimental technique that provides good selectivity. Where conditions are chosen to provide high efficiency, two narrow and distinct peaks are obtained (high resolution). When the conditions provide only low efficiency, the peaks are broadened and overlap more (lower resolution). (**B**) Chromatogram obtained using experimental technique that provides bad selectivity. Even when high-efficiency conditions are used, there is lower resolution than in (A), and the resolution deteriorates further with low-efficiency conditions.

passed through the column at peak maximum) and W_h is the peak width at half the peak height.

$$N = 5.54\left(\frac{V_R}{W_h}\right)^2$$

Equation 8.1.4

Efficiency (N) may be expressed as the number of theoretical plates for the column under specific experimental conditions (see *UNIT 8.3*). Efficiency is also frequently defined as the number of plates per meter of chromatographic bed, or in terms of H, the height equivalent to a theoretical plate. H is simply the column length (L) divided by the efficiency factor (N); i.e., $H = L/N$.

The main cause of zone broadening (i.e., loss of efficiency) in a chromatographic bed is diffusion. Diffusion perpendicular to the flow is restricted by the walls of the column. Therefore, longitudinal diffusion is the primary fac-

PCU IASI / CENTRAL UNIVERSITY LIBRARY

tor contributing to zone broadening. While a protein is adsorbed to the medium, little or no diffusion takes place, but once the protein is unbound, diffusion begins. The amount of diffusion that occurs is proportional to the time required for the material to emerge from the system. Loss of efficiency resulting from diffusion is minimized if the distances available for diffusion in the mobile phase, gel beads, and system are minimized. In practice, efficiency increases with increasing uniformity in particle size and with decreasing bead size. Good experimental technique is required for high efficiency. Unevenly packed chromatography beds and trapped air will lead to channeling, zone broadening, and consequent loss of resolution. Loss of efficiency also stems from system effects such as dead volumes in the system, poor mixing during gradient formation, and pulsations in flow.

Selectivity

Selectivity of a chromatographic medium defines the ability of that medium to separate peaks (i.e., it is a measure of the distance between two peaks in a chromatogram; see Fig. 8.1.5). The selectivity factor (α) can be calculated from a chromatogram using Equation 8.1.5, where k_2 is the capacity factor for the second peak (see discussion of Capacity), k_1 is the capacity factor for the first peak, V_{R2} is the elution volume of the second peak, V_{R1} is the elution volume of the first peak (see Fig. 8.1.2), and V_m is the volume of the mobile phase.

$$\alpha = \frac{k_2}{k_1} = \frac{V_{R2} - V_m}{V_{R1} - V_m} \approx \frac{V_{R2}}{V_{R1}}$$

Equation 8.1.5

Selectivity is more important than efficiency (N) in determining resolution because R_s is directly proportional to selectivity, but is proportional only to the square root of efficiency (see Fig. 8.1.5 along with Equation 8.1.1 and Equation 8.1.2 for R_s in the discussion of Resolution). Hence, a four-fold increase in efficiency is required to double resolution, as compared with a two-fold increase in selectivity. In practice, selectivity depends partly on the chromatographic technique employed but can usually be controlled by manipulating experimental conditions, such as the pH and ionic strength of the mobile phase. Because this can be done easily and predictably, selectivity is the factor that is exploited to achieve maximum resolution in column chromatography rather than efficiency, which is fixed by the particle size and uniformity of the medium selected.

LITERATURE CITED

Corran, P.H. 1989. Reversed-phase chromatography of proteins. *In* HPLC of Macromolecules: A Practical Approach (R.W.A. Oliver, ed.) pp. 127-156. IRL Press, Oxford.

Giddings, J.C. and Keller, R.A. (eds.) 1965. Dynamics of Chromatography, Part 1: Principles and Theory. Marcel Dekker, New York.

Janson, J.-C. and Ryden, L. (eds.) 1989. Protein Purification: Principles, High Resolution Methods and Applications. VCH Publishers, New York.

Contributed by Alan Williams
Pharmacia Biotech
Piscataway, New Jersey

B.C.U. „M. Kogălniceanu" IAŞI
CHIMIE

BCU IAŞI / CENTRAL UNIVERSITARĂ LIBRARY

Ion-Exchange Chromatography

Ion-exchange chromatography separates biomolecules on the basis of charge characteristics. Charged groups on the surface of a protein interact with oppositely charged groups immobilized on the ion-exchange medium. As illustrated in Figure 8.2.1, the charge of a protein depends on the pH of its environment (the operating pH). The pH at which the net charge of a protein is zero (i.e., where the number of positive charges equals the number of negative charges) is known as the isoelectric point (pI). When the operating pH is greater than the pI, the protein will have a net negative charge, and should bind to anion-exchange media, which are positively charged. When the operating pH is less than the pI, the protein will have a net positive charge, and should bind to cation-exchange media, which are negatively charged.

The Strategic Planning section outlines the basic steps in planning and carrying out ion-exchange chromatography to separate proteins. Basic Protocol 1 describes batch adsorption of protein to an ion-exchange medium followed by elution using a step gradient of increasing salt concentration. This technique accommodates a wide range of sample volumes and is most often used in the capture stage of protein purification (*UNIT 8.1*). Batch techniques have minimal system requirements. The Alternate Protocol describes use of a step gradient of increasing pH in batch purification. Basic Protocol 2 describes adsorption of protein to an ion-exchange medium in a column, followed by elution with a linear gradient. Basic Protocol 2 provides a higher resolution than Basic Protocol 1, and is therefore used in the intermediate purification and polishing stages of protein separation (*UNIT 8.1*). Support Protocol 1 describes a pilot experiment to determine initial conditions for batch or column chromatography (i.e., pH required for binding, change in pH or salt concentration required for elution, and available capacity of a medium). Support Protocol 2 describes a means of calculating the dynamic capacity (*UNIT 8.1*) of an ion-exchange column, Support Protocol 3 describes methods for producing continuous gradients of pH and salt concentration to elute proteins from ion-exchange columns, Support Protocol 4

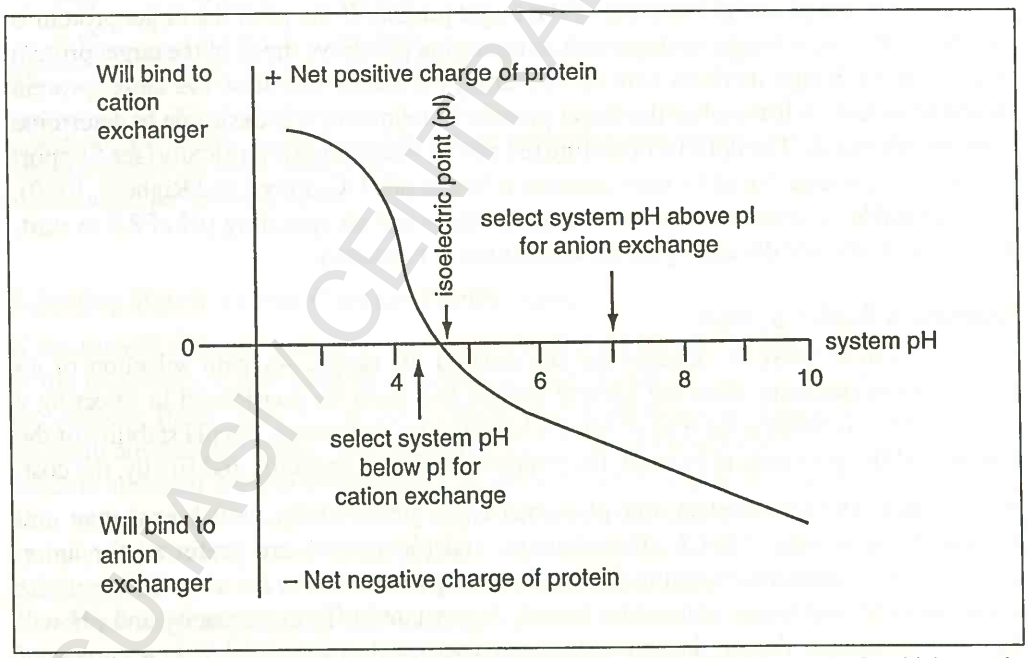


Figure 8.2.1 Net charge of a protein as a function of pH, showing the pH ranges in which protein is bound to anion or cation exchangers. The pH range over which the protein is stable may be only a small fraction of the binding range; this must also be taken into consideration when choosing an ion-exchange medium.

Copyright © 1995 by John Wiley & Sons, Inc.

UCLA CENTRAL UNIVERSITY LIBRARY

describes regeneration of used ion-exchange media, and Support Protocol 5 details storage of ion-exchange media.

STRATEGIC PLANNING

There are a number of basic steps in developing an ion-exchange method for protein purification. First, an appropriate ion-exchange medium must be selected, as well as the optimal operating pH and buffer system for the medium and sample (see following discussions on Selecting an Ion-Exchange Medium and Selecting a Buffer System). It is then necessary to decide whether batch or column chromatography is appropriate given the physical characteristics of the sample, the sample volume to be used, and the availability of suitable equipment (see following discussion on Selecting Batch Versus Column Purification). Next, pilot experiments are conducted to determine conditions for binding and eluting the protein (see Support Protocol 1). It is also necessary to make sure that the capacity of the medium is sufficient to isolate the desired quantity of protein (see Support Protocol 1 and Support Protocol 2). When these initial conditions have been determined, the medium and sample are prepared, the sample is bound to the medium, unbound sample components are washed away, and bound sample components are selectively eluted and collected (see Basic Protocol 1, Basic Protocol 2, or Alternate Protocol). The medium may then be cleaned and regenerated (see Support Protocol 5). Finally, results are evaluated and conditions optimized if necessary (see Critical Parameters).

Selecting an Ion-Exchange Medium

A wide variety of ion-exchange media are commercially available, but no miracle medium exists that is best for every protein purification. Criteria for selecting an ion-exchange medium include the specific requirements of the application, the pI and molecular size of the sample components (i.e., target protein and contaminants), and the available equipment (e.g., pumps and columns).

It is necessary to begin by selecting either anion or cation exchange. This requires knowledge of the pI and pH stability of the target protein. If the pI of the target protein is known, an anion-exchange medium with an operating pH above the pI of the target protein or a cation-exchange medium with an operating pH below the pI of the target protein should be selected. If the pI of the target protein is unknown, it is desirable to determine it before beginning. The optimal operating pH can be determined empirically (see Support Protocol 1). Because the pI for most proteins is below pH 7 (Gianazzi and Righetti, 1980), it is reasonable to select an anion-exchange medium and an operating pH of 8.5 to start, then evaluate the results and optimize conditions as necessary.

Selecting a Buffer System

A buffer system must be selected for the desired pH range. As with selection of an ion-exchange medium, there are several factors that must be considered in selecting a buffer system, including the type of ion exchange to be performed, the pH stability of the sample and the pH range to be used, the required buffering capacity, and finally, the cost.

Anionic buffers (e.g., acetate and phosphate) are preferred for cation exchange and cationic buffers (e.g., Tris·Cl, ethanolamine, and piperazine) are preferred for anion exchange. It is important to ensure that the buffering ion will have the same charge as the ion exchanger, and hence will not be bound. A constant buffering capacity and pH will thus be maintained during the ion-exchange experiment. Table 8.2.1 lists a variety of buffers useful for anion and cation exchange. Additives to be included in the buffers (e.g., detergents or protease inhibitors) should also carry the same charge as the ion-exchange medium, to preclude binding.

RCU IASI CENTRAL UNIVERSITY LIBRARY

Table 8.2.1 Buffers for Ion-Exchange Chromatography[a]

pK_a (25°C)	pH range	Buffer[b]	Working concentration (mM)	Temperature factor[c]
Anion exchange				
4.75	4.5-5.0	*N*-methylpiperazine	20	–0.015
5.68	5.0-6.0	Piperazine	20	–0.015
5.96	5.5-6.0	L-histidine	20	
6.46	5.8-6.4	bis-Tris	20	–0.017
6.80	6.4-7.3	bis-Tris propane	20	
7.76	7.3-7.7	Triethanolamine	20	–0.02
8.06	7.6-8.5	Tris·Cl	20	–0.028
8.52	8.0-8.5	*N*-methyldiethanolamine	50	–0.028
8.88	8.4-8.8	Diethanolamine	20 (pH 8.4)	–0.025
8.64	8.5-9.0	1,3-diaminopropane	20	–0.031
9.50	9.0-9.5	Ethanolamine	20	–0.029
9.73	9.5-9.8	Piperazine	20	–0.026
10.47	9.8-10.3	1,3-diaminopropane	20	–0.026
11.12	10.6-11.6	Piperadine	20	–0.031
Cation exchange				
2.00	1.5-2.5	Maleic acid	20	
2.88	2.4-3.4	Malonic acid	20	
3.13	2.6-3.6	Citric acid	20	–0.0024
3.81	3.6-4.3	Lactic acid	50	
3.75	3.8-4.3	Formic acid	50	+0.0002
4.21	4.3-4.8	Butanedioic acid	50	–0.0018
4.76	4.8-5.2	Acetic acid	50	+0.0002
5.68	5.0-6.0	Malonic acid	50	
7.20	6.7-7.6	Phosphate	50	–0.0028
7.55	7.6-8.2	HEPES	50	–0.0140
8.35	8.2-8.7	BICINE	50	–0.0180

[a]Information from Pharmacia Biotech, 1991.

[b]Abbreviations: BICINE, *N,N-bis*[2-hydroxyethylglycine; bis-Tris, bis[2-hydroxyethyl]iminotris[hydroxymethyl]methane; bis-Tris propane, 1,3-bis[tris(hydroxymethyl)methylamino]-propane; HEPES, *N*-[2-hydroxyethyl]piperazine-*N'*-[2-ethanesulfonic acid].

[c]Change in pK_a per °C (i.e., dpK_a/dT).

Selecting Batch Versus Column Purification

If the sample volume is large in relation to the size of the pumps and columns available in the laboratory, batch adsorption techniques (see Basic Protocol 1) are appropriate for the capture stage (*UNIT 8.1*) of purification. Batch methods, in which the sample and medium are directly mixed without use of a column, are employed to reduce the sample volume and total protein content prior to column chromatography. Batch techniques are also very useful for capture of the target protein from extremely crude samples (e.g., samples containing particulates and/or aggregates). In batch methods the flow and packing characteristics of the medium are of minor importance, but economy and capacity are of primary importance. Ion-exchange media suitable for batch purification at the capture stage of protein purification are presented in Table 8.2.2.

High resolution is required in the intermediate purification and polishing stages of protein purification (*UNIT 8.1*) and this can only be achieved using column chromatography.

NOT FOR CENTRAL UNIVERSITY LIBRARY

Table 8.2.2 Ion-Exchange Media for Batch Adsorption

Medium	Type	Supplier[a]
Dextran bead matrix		
QAE Sephadex A-25	Strong anion	Pharmacia Biotech
QAE Sephadex A-50	Strong anion	Pharmacia Biotech
SP Sephadex C-25	Strong cation	Pharmacia Biotech
SP Sephadex C-50	Strong cation	Pharmacia Biotech
Microgranular cellulose matrix		
QA-52	Strong anion	Whatman
QA-53	Strong anion	Whatman
SE-52	Strong cation	Whatman
SE-53	Strong cation	Whatman
Polymer-coated silica matrix		
Accell Plus QMA	Strong anion	Waters
Accell Plus CM	Weak cation	Waters
Agarose bead matrix		
DEAE Bio-Gel A	Weak anion	Bio-Rad
CM Bio-Gel A	Weak cation	Bio-Rad

[a]For addresses and telephone numbers of suppliers, see *SUPPLIERS APPENDIX*.

However, higher-than-necessary resolution is frequently traded off for greater throughput (i.e., amount of material processed in a defined time). The packing and flow characteristics of the medium are of primary importance when using column chromatography. Columns packed with smaller beads usually require higher operating pressures than columns packed with larger beads. However, smaller beads offer higher resolution, resulting from the increased efficiency that comes with decreasing bead size (*UNIT 8.1*). Larger beads are desirable for applications requiring higher throughput. A medium with the smallest bead size that will provide the necessary throughput (and whose operating pressure can be accommodated by available equipment) should be selected. Table 8.2.3 lists a variety of media suitable for column purification at various stages in the purification process. See *SUPPLIERS APPENDIX* for contact information of suppliers of these media, all of whom provide technical support for selection and use of specific media.

Capacity of Ion-Exchange Media

The amount of ion-exchange medium required for a particular application will be determined by the protein capacity of the medium (i.e., the amount of protein that can be bound to the matrix) at the chosen pH. The amount of medium required will in turn determine the size of column required. The capacity of an ion-exchange medium is dependent on a variety of factors, including charge and molecular size of the components in the sample and experimental conditions employed (also see *UNIT 8.1*).

The number of charged groups on a specified amount of matrix that can participate in ion exchange is termed **ionic** or **total** capacity. However, ion-exchange media are porous and the ion-exchange groups are usually distributed throughout the bead. The amount of protein that will bind is dependent on the ability of the sample components to diffuse into the bead, and this ability is related to the molecular size of the components and the pore structure of the matrix (see *UNIT 8.3*). In addition, large molecules such as proteins can bind to multiple ion-exchange groups (or sterically hinder other molecules from binding) depending on the shape and surface-charge distribution of the protein, as well as the spatial

BCU IAS CENTRAL UNIVERSITY LIBRARY

Table 8.2.3 Ion-Exchange Media for Column Purification[a]

Medium[b]	Matrix	Avg. bead diameter (μm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
Pharmacia Biotech								
Mono Q	PS/DVB	10	I/P	PP/BU	320	–	65	4.5 mg/ml BSA; 50 mM Tris·Cl, pH 8.0; 100 cm/hr; 0.5 × 5–cm column
Mono S	PS/DVB	10	I/P	PP	160	–	75	Human IgG; 100 mM sodium acetate, pH 5.0; 300 cm/hr; 0.5 × 5–cm column
SOURCE 15Q	PS/DVB	15	I/P	PP/BU	ND	–	54	4.5 mg/ml BSA; 50 mM Tris·Cl, pH 8.0; 100 cm/hr; 0.5 × 5–cm column
SOURCE 15S	PS/DVB	15	I/P	PP/BU	ND	–	85	55 mg/ml lysozyme; 20 mM phosphate, pH 6.8; 1 ml/min; 0.63 × 3–cm column
Q Sepharose High Performance	Agarose	34	C/I/P	PP/BU	170	–	108	4.5 mg/ml BSA; 5 mM Tris·Cl, pH 8.0; 100 cm/hr; 0.5 × 5–cm column
SP Sepharose High Performance	Agarose	34	C/I/P	PP/BU	170	–	90	RNase; 100 mM sodium acetate, pH 5.0; 150 cm/hr; 0.5 × 5–cm column
Q Sepharose Fast Flow	Agarose	90	C/I	BU	215	–	120	Human serum albumin; 50 mM Tris·Cl, pH 8.3
SP Sepharose Fast Flow	Agarose	90	C/I	BU	215	–	50	Human IgG; 100 mM sodium acetate, pH 5.0; 300 cm/hr; 0.5 × 5–cm column
SP Sepharose Big Beads	Agarose	200	C	BU	210	–	80	2 mg/ml BSA; 100 mM sodium acetate, pH 5.0; 300 cm/hr
STREAMLINE DEAE	Agarose	200	B/C	BU	205	–	55	BSA; 50 mM Tris·Cl, pH 7.5; 300 cm/hr; 5-cm–i.d. expanded-bed column
STREAMLINE SP	Agarose	200	B/C	BU	170	–	70	Lysozyme; 50 mM phosphate, pH 7.5, 300 cm/hr; 5-cm–i.d. expanded-bed column
Hi-Trap Q[c]								
HiTrap S[d]								
Bio-Rad								
Macro-Prep high Q	PM	10	I/P	PP/BU	115	–	20	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.3; 100 cm/hr; 1.5 × 10–cm column
Macro-Prep high S	PM	10	I/P	PP/BU	127	–	50	5 mg/ml human IgG; 20 mM sodium acetate, pH 5.0; 100 cm/hr; 1.5 × 10–cm column
Macro-Prep high Q	PM	50	B/C/I	PP/BU	400	–	40	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.3; 100 cm/hr; 1.5 × 10–cm column

continued

BCU HASI / CENTRAL UNIVERSITY LIBRARY

Table 8.2.3 Ion-Exchange Media for Column Purification[a], continued

Medium[b]	Matrix	Avg. bead diameter (μm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
Macro-Prep high S	PM	50	B/C/I	P/B	160	–	60	5 mg/ml human IgG; 20 mM sodium acetate, pH 5.0; 100 cm/hr; 1.5 × 10–cm column
Econo-Pac high Q[e]								
Econo-Pac high S[f]								
TosoHaas								
Toyopearl QAE-550C	PM	100	C/I	PP/BU	330	70[g]	ND	–
Toyopearl SP-550C	PM	100	C/I	PP/BU	160	110[h]	ND	–
Toyopearl SUPER Q-650C	PM	100	C/I	PP/BU	280	129[g]	ND	–
Toyopearl SUPER Q-650M	PM	65	C/I/P	PP/BU	240	143[g]	ND	–
Toyopearl SUPER Q-650S	PM	35	I/P	PP/BU	240	126[g]	ND	–
Toyopearl SP-650C	PM	100	C/I	PP/BU	150	45[h]	ND	–
Toyopearl SP-650M	PM	65	C/I/P	PP/BU	150	50[h]	ND	–
Toyopearl SP-650S	PM	35	I/P	PP/BU	150	50[h]	ND	–
Waters								
Protein-Pak Q HR	PM	8	P	PP	200	–	65	BSA; 20 mM Tris·Cl, pH 8.2
Protein-Pak SP HR	PM	8	P	PP	225	–	40	Cytochrome c; 25 mM MES, pH 5.0
Protein-Pak Q HR	PM	15	I/P	PP/BU	200	–	75	BSA; 20 mM Tris·Cl, pH 8.2
Protein-Pak SP HR	PM	15	I/P	PP/BU	225	–	40	Cytochrome c; 25 mM MES, pH 5.0
Protein-Pak Q HR	PM	40	B/C/I	BU	200	–	60	BSA; 20 mM Tris·Cl, pH 8.2
Protein-Pak SP HR	PM	40	B/C/I	BU	225	–	20	Cytochrome c; 25 mM MES, pH 5.0
BioSepra								
Q-Hyper D	CP	10	P	PP	250	–	90	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	10	P	PP	110	–	105	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough

continued

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 8.2.3 Ion-Exchange Media for Column Purification[a], continued

Medium[b]	Matrix	Avg. bead diameter (μm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
Q-Hyper D	CP	20	I/P	PP/BU	250	–	90	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	20	I/P	PP/BU	100	–	105	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough
Q-Hyper D	CP	35	I/P	PP/BU	250	–	40	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	35	I/P	PP/BU	100	–	100	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough
Q-Hyper D	CP	60	C/I	BU	200	–	100	5 mg/ml BSA; 50 mM Tris·Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	60	C/I	BU	100	–	100	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough

[a]Abbreviations: B, batch; BSA, bovine serum albumin; BU, bulk; C, capture; CP, composite polymer, I, intermediate purification; i.d., inner diameter; MES, 2-[*N*-morpholino]ethanesulfonic acid; ND, not determined; P, polishing; PM, polymethacrylate; PP, prepacked; PS/DVB, polystyrene/divinylbenzene.

[b]For addresses and telephone numbers of suppliers, see *SUPPLIERS APPENDIX*.

[c] Prepacked column suitable for use with syringe. For medium specification, see Q Sepharose High Performance.

[d] Prepacked column suitable for use with syringe. For medium specification, see SP Sepharose High Performance.

[e] Prepacked column suitable for use with syringe. For medium specification, see Macro-Prep high Q.

[f] Prepacked column suitable for use with syringe. For medium specification, see Macro-Prep high S.

[g]Conditions for determining available capacity: 5 mg/ml BSA; 50 mM Tris·Cl, pH 8.7.

[h]Conditions for determining available capacity: 5 mg/ml lysozyme; 20 mM phosphate, pH 6.0.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

orientation of the exchange ligands on the matrix. These factors are taken into account in the available and dynamic capacity of a matrix or column.

Available capacity is defined as the amount of a specific protein that will bind to an ion-exchange medium at a defined pH, salt concentration, and sample concentration in a batch purification process. When medium is packed in a column, the amount of a specific protein that will bind is dependent on many of these factors, but also on the column dimensions and flow rate. The capacity of a column operating under defined conditions is termed the **dynamic capacity**. The dynamic capacity is typically less than the available capacity for the same sample under the same operating conditions.

Ion-exchange media in which the charge state of the ligand is sensitive to the operating pH are termed weak ion exchangers. The usable pH range of weak ion exchangers is narrower than that of strong ion-exchange media, in which the charge on the ligand used for exchange is constant (i.e., insensitive to pH). Ion-exchange media with DEAE (diethylaminoethyl) or CM (carboxymethyl) groups as the ligand for exchange are considered weak ion exchangers. Because the ionic capacity of weak ion-exchange media is variable, strong ion-exchange media are recommended for pilot experiments (see Support Protocol 1).

When selecting ion-exchange media for batch or column purifications, the amount of medium required should be estimated from the manufacturer-supplied capacity data for a protein similar in size to the target protein to be purified. Be aware of what type of capacity (i.e., ionic, available, or dynamic) is being reported and the experimental conditions under which the reported capacity was determined (see Table 8.2.3). A general recommendation is to begin pilot experiments at ~10% of the published capacity data, then determine the capacity for your system.

Preparing Solutions

All aqueous solutions for use in ion-exchange chromatography should be prepared with distilled or deionized water (Milli-Q purified or equivalent). Solutions and samples for column chromatography should be filtered, degassed, and equilibrated to the appropriate temperature prior to use. Special attention must be given to any additives to be included in the buffer to be sure that they will not bind to the column (see preceding discussion on Selecting a Buffer System). If the chosen additives increase the viscosity of the mobile phase, the flow rate should be decreased.

Chromatography Systems

Column techniques typically result in higher resolution than batch techniques, and should be used during the intermediate purification and polishing stages (*UNIT 8.1*) of a purification procedure. However, to achieve high resolution, an appropriate chromatography system is required. A chromatography system may be described simply by the functional operations required to perform a chromatographic separation. The minimum functions required (Fig. 8.2.2) include a buffer delivery system capable of gradient formation, a mechanism for introducing sample into the system, the column itself (properly packed with the appropriate ion exchanger), an in-line UV monitor suitable for use at 280 nm, a chart recorder, and a fraction collector to preserve the separation. Additional unit operations may include system controllers for automated operation as well as data-acquisition systems. The equipment required to perform each of these unit operations is ultimately determined by the chromatography medium and column selected. Specifically, the operating pressure and flow characteristics of the packed column will be the most important factors for selecting equipment to fit the column, or for selecting columns to fit existing equipment.

RCU IASRI CENTRAL UNIVERSITY LIBRARY

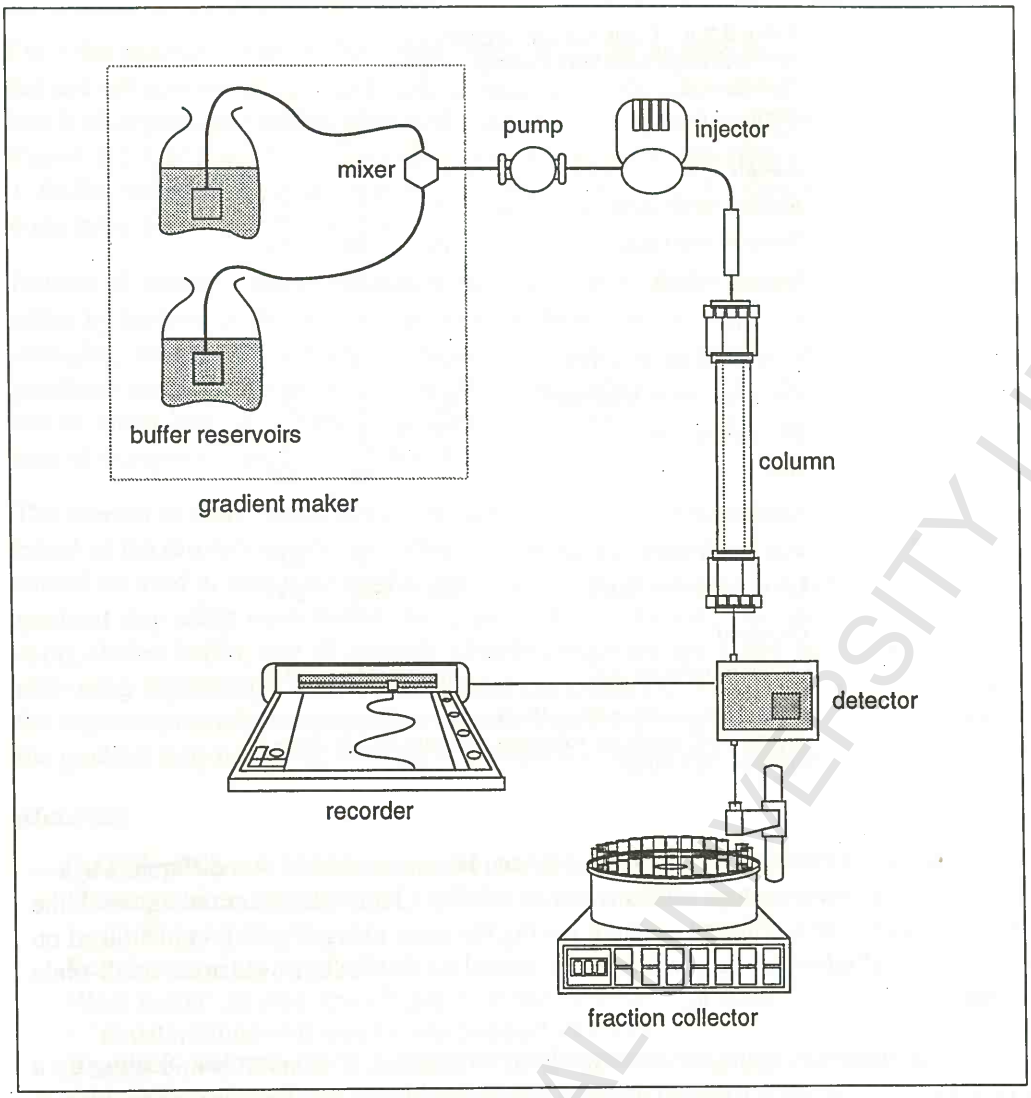


Figure 8.2.2 Liquid column chromatography system with gradient maker.

It is highly recommended that prepacked columns be used and that the manufacturer be consulted prior to purchase as to equipment requirements for proper operation. This will save time and money and increase the probability of success. If columns are to be packed in the laboratory, the manufacturer's packing instructions for the medium should be followed and column efficiency and function should be tested with representative standards prior to use with the actual sample. There is no single packing methodology that is optimal for all types of media. Examples of packing techniques suitable for use with ion exchangers are included in *UNIT 8.4*. Table 8.2.3 lists a variety of ion-exchange media available either in prepacked columns and/or in bulk for packing in the laboratory. Several manufacturers offer inexpensive disposable ion-exchange columns, suitable for many applications, that can be operated using only a syringe.

Scale-Up Conditions

Ion-exchange separations are typically developed and optimized on a small scale before the entire sample is committed. It is thus important to choose an ion exchanger that will allow simple and convenient scale-up, so that methods established on a small column can be applied more or less directly to a larger one.

B.C.U. „M. I... SCU" IAŞI
CHI IE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Table 8.2.4 Scale-up Calculations for Ion-Exchange Column Chromatography

Scale	1×	10×	100×
Sample			
Sample concentration (mg/ml)	1	1	1
Sample load (mg)	10	100	1000
Sample volume (ml)	10	100	1000
Column			
Column diameter (cm)	0.5	1.6	5
Cross-sectional area (cm^2)	0.2	2	20
Bed height (cm)	5	5	5
Bed volume (ml)	1	10	100
Flow Rate			
Volumetric flow (ml/min)[a]	2	20	200
Linear velocity (cm/hr)	300	300	300
Gradient			
Gradient volume (ml)	20	200	2000

[a]Linear flow (cm/hr) multipled by the cross-sectional area of the column (cm^2) equals the volumetric flow rate (ml/hr); cm/hr × cm^2 = cm^3/hr = ml/hr.

Scaling up in ion-exchange chromatography can be approached in two different ways—either the same ion-exchange medium can be used in a larger-diameter column with the same bed height, or a different medium having the same charged group immobilized on a matrix with a higher throughput can be substituted for the medium used in the small-scale separation.

The former approach, using the same medium throughout, is the simplest. Scaling up a separation with the same medium and bed height used in the small-scale optimization is achieved by employing simple scale factors to adapt the flow, gradient volume, and sample load to the increased volume of a larger-diameter column. The scale factor for the required increase in flow rate is the ratio of the cross-sectional area of the larger column to that of the smaller column (i.e., the factor by which the total flow of solution through the column must be increased so that the same linear velocity used in the small-scale experiment is maintained). The increase in gradient volume and sample loading are directly proportional to the increase in column volume. Table 8.2.4 shows an example of changes that must be made to various parameters for a 10- and 100-fold increase in scale.

BASIC PROTOCOL 1

BATCH ADSORPTION AND STEP-GRADIENT ELUTION WITH INCREASING SALT CONCENTRATION

This protocol describes adsorption of a protein to an ion-exchange medium by direct mixing of the protein-containing sample and the medium (in this case, an anion-exchange gel). This batch purification is in contrast to column purification (see Basic Protocol 2), in which the gel is packed in a column and the sample is passed through it. Batch adsorption is particularly useful where large volumes of sample are involved, which is usually the case in the capture stage of protein purification (*UNIT 8.1*). The equipment used here is sufficient for sample volumes up to 1.5 liter. Larger samples may be divided and processed with parallel apparatuses, or processed sequentially on a single apparatus. Different-sized funnels and side-arm flasks may be used when appropriate.

RCU IASE/CENTRAL UNIVERSITY LIBRARY

Once the starting conditions have been determined (see Support Protocol 1), the range of pH and salt concentration where binding and elution occur is well defined. Conditions for batch adsorption and elution used in this procedure are based on the results, presented in Figure 8.2.3, of a test tube pilot experiment performed as described in Support Protocol 1. Buffer volumes and equipment were chosen on the basis of a sample volume of 1 liter containing 50 to 1000 mg of total protein.

Elution of sample components adsorbed to an ion-exchange medium can be achieved either by increasing the salt concentration of the mobile phase, as in this protocol, or by changing its pH (see Alternate Protocol). In batch techniques, discontinuous (step) gradients are used for elution. Column chromatography (see Basic Protocol 2) allows the use of either step or continuous gradients. Formation of continuous gradients requires special equipment (see Basic Protocol 2 and Support Protocol 3).

The number of incremental steps to include in a step gradient must be judged empirically based on the resolution achieved. Ideally (as in the procedure described here), three steps should be used in which all weakly bound materials are desorbed and eluted in the first gradient step using wash buffer, the target protein is eluted in the second gradient step using elution buffer, and all strongly adsorbed materials are eluted in the third gradient step using regeneration buffer. A minimum increment of 50 mM NaCl is recommended for experiments using step-gradient elution with increasing salt concentration. Steps in the gradient may be added, omitted, or modified as determined empirically.

Materials

QAE Sephadex A-25 (Pharmacia Biotech) or equivalent anion-exchange gel
Binding buffer: 20 mM Tris·Cl, pH 7.5 (or other buffer as determined empirically; see Support Protocol 1)
Protein sample to be purified
Wash buffer: 20 mM Tris·Cl (pH 7.5)/100 mM NaCl (or other buffer/salt solution as determined empirically; see Support Protocol 1)
Elution buffer: 20 mM Tris·Cl (pH 7.5)/350 mM NaCl (or other buffer/salt solution as determined empirically; see Support Protocol 1)
Regeneration buffer: 20 mM Tris·Cl (pH 7.5)/2 M NaCl (also see Support Protocol 5)

Boiling water bath (optional)
500-ml sintered-glass filter funnel, medium porosity
Three 2000-ml side-arm flasks
Conductivity meter

Swell and equilibrate ion-exchange gel

1. Add 10 g QAE Sephadex A-25 to 1 liter binding buffer and swell 2 days at room temperature or 2 hr in a boiling water bath.

   The volume of swollen gel will be ~70 ml.

2. Mount 500-ml sintered-glass filter funnel on 2000-ml side-arm flask and apply suction.

   Centrifugation (~5000 × g, 1 min) may be used instead of filtration to collect the gel, and may be preferred for processing small sample volumes. It is also possible to simply allow the gel to settle under gravity, then decant or aspirate the supernatant; this may be preferable for very large-scale operations.

3. Gently swirl swollen gel to resuspend. Pour gel slurry into funnel as fast as the fluid level in funnel permits, allowing buffer to collect in flask. Continue until all gel has been collected. Release suction after all buffer has been removed.

EDINBURGH CENTRAL UNIVERSITY LIBRARY

IMPORTANT NOTE: *Never use a magnetic stir bar to resuspend chromatography medium. This can damage beads by grinding them against the bottom of the vessel, resulting in fine particles that can slow the filtration process.*

4. Add 200 ml binding buffer to funnel and resuspend gel using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer.

5. Repeat step 4 at least three times to ensure equilibration.

   The gel is considered equilibrated when the pH and salt concentration of the eluant are the same as those of the binding buffer. If the procedures here are followed, there is little need to measure pH and salt concentration of the eluant because the gel is swollen and washed in binding buffer and the chance of error is very small. The worst-case scenario is that the target protein will not bind (see Troubleshooting). The repeated washings are more for the removal of fines than for adjustment of pH or salt concentration.

6. Add enough binding buffer to cover gel (~100 ml, to produce an ~50% slurry) and allow gel to stand in funnel until sample has been prepared.

Adsorb sample to gel

7. Adjust pH and salt concentration of protein sample to be purified to initial optimal values.

   For a pilot experiment to determine optimal initial values for pH and sale concentration, see Support Protocol 1.

   Samples prepared by salt precipitation may require desalting prior to ion exchange (see UNIT 8.3). The salt concentration of a sample can be reduced by adding distilled water until a desired salt concentration is achieved, as measured by a conductivity meter.

8. Combine gel and sample in 2000-ml beaker or wide-mouth flask.

   The ~50% gel slurry is swirled and poured from the funnel; transfer may be aided with a rubber policeman, spoon, stir rod, or wash bottle containing binding buffer.

9. Gently mix sample and gel by swirling every 15 min or shake on a platform shaker at sufficient rate to maintain gel in suspension. Allow 1 to 2 hr for binding at room temperature or 3 to 4 hr at 4°C.

   IMPORTANT NOTE: *Excessive shaking may result in foaming and possible denaturation of the target protein.*

10. Collect gel by filtration as in step 3. Save filtrate for assay.

    This filtrate is saved in case the target protein did not bind. All filtrates should be kept until the fraction containing the target protein has been identified. Storage conditions depend on the stability of the target protein; as a general rule they should be refrigerated to minimize microbial growth. Bacteriostatic agents may be added for prolonged storage.

11. Add 100 ml binding buffer to funnel and resuspend gel using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Repeat three or four times, pooling all washings with the filtrate from step 10.

    The gel containing the bound sample may be packed in a column for subsequent elution steps (see Basic Protocol 2).

Elute with step gradient of increasing salt concentration

12. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

13. Add an equal volume of wash buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Save filtrate.

The wash buffer is the first step of the gradient. The filtrate will contain sample components that were weakly adsorbed to the gel (see annotation to step 10).

14. Measure absorbance of filtrate at 280 nm using wash buffer as blank.

15. Repeat steps 13 and 14 until the last filtrate shows no significant absorbance at 280 nm. Pool and save the filtrates.

16. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

17. Add an equal volume of elution buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand 5 min, then apply suction to remove buffer. Save filtrate.

The elution buffer is the second step of the gradient. The filtrate will contain the sample components that were more strongly adsorbed to the gel and should include the target protein.

18. Measure the absorbance of the filtrate at 280 nm using elution buffer as a blank.

19. Repeat steps 17 and 18 until the last filtrate shows no significant absorbance at 280 nm. Pool and save the filtrates for assay and subsequent purification of the target protein.

If gel is to be reused, continue with remaining steps. Otherwise, discard the gel.

Regenerate ion-exchange medium for reuse

20. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

21. Add an equal volume of regeneration buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Repeat five times and discard filtrates.

The regeneration buffer removes strongly bound materials from the gel. See Support Protocol 5 for additional considerations in regenerating ion-exchange media.

22. Repeat step 11 for a total of five washes to reequilibrate the gel with binding buffer.

23. Check the pH and conductivity of the last filtrate to ensure that the gel is properly equilibrated.

See Support Protocol 5 for information on storing ion-exchange media.

pH-BASED STEP-GRADIENT ELUTION

ALTERNATE PROTOCOL

The net charge of a protein is pH dependent. Therefore, altering the pH of the mobile phase to make it closer to the pI of a protein can change its net charge, causing it to desorb and elute from an ion exchanger. Continuous pH gradients are very difficult to produce at constant salt concentration because charge characteristics of sample components, buffering ions, and ion-exchange media depend on the pH of the system. pH-based step gradients are much easier to produce and more reproducible.

The first basic protocol can be modified for pH-gradient elution by selecting a buffer at a pH suitable for elution, based on results obtained in a pilot experiment (see Support Protocol 1). This elution buffer is then used in steps 17 to 19 of the first basic protocol. Incubation times and buffer volumes should be increased by 20% when using pH elution.

BU ARI CENTRAL UNIVERSITY LIBRARY

When scouting methods for ion exchange, it is best to start with changes in salt concentration (see Basic Protocol 1) as this technique is simpler and more reproducible. Elution with pH change is only recommended when the ion-exchange behavior of the sample is well known and the resolution required cannot be attained with a change in salt concentration. pH elution may also be appropriate in cases where it provides a superior ionic/pH environment for loading in the next chromatographic step. A combination of increasing salt concentration and pH change may also be used for elution.

BASIC PROTOCOL 2

COLUMN CHROMATOGRAPHY WITH LINEAR GRADIENT ELUTION

Column chromatography is generally preferred to batch chromatography (see Basic Protocol 1), especially where high resolution is required, although it accommodates smaller sample volumes. Column chromatography requires specialized equipment to achieve high resolution, including a column containing the medium and a fluid-delivery system appropriate for the pressure and flow rate at which the column is to be operated. It is the technique preferred for the intermediate purification and polishing stages of protein purification (*UNIT 8.1*). Also in contrast to the first basic protocol, this technique uses a linear gradient for elution. Although more difficult to produce than step gradients (see Basic Protocol 1), linear gradients lead to better resolution (see Support Protocol 3).

The example presented here uses a Resource Q anion-exchange column from Pharmacia Biotech, which may be used with either HPLC (high-performance liquid chromatography) or FPLC (fast protein liquid chromatography) systems. Optimal conditions for other columns should be obtained from the manufacturer. Procedures mentioned in the protocol steps should be carried out according to the manufacturer's instructions for the particular chromatography system being used.

NOTE: All buffers, media, and other system components should be filtered, degassed, and equilibrated to the same temperature before use.

Materials

Liquid chromatography system (FPLC or HPLC)
Elution buffer: binding buffer (see Basic Protocol 1 and Support Protocol 1) containing 1 M NaCl
Binding buffer (see Basic Protocol 1 and Support Protocol 1)
Resource Q chromatography column (1-ml packed bed volume; Pharmacia Biotech)
Protein sample to be purified

Conductivity meter
0.22-µm filter

Prepare chromatography system

1. Set up the liquid chromatography system according to manufacturer's instructions, without the column in-line.

2. Test system performance by running a blank gradient ranging from 0% to 100% elution buffer in 20 ml at a constant flow rate of 5 ml/min.

   See Figure 8.2.2 for a diagram of a typical column chromatography system. The gradient is formed by placing binding and elution buffer in the appropriate buffer reservoirs (see Support Protocol 3 for detailed discussion of continuous gradient formation).

BCU IASI CENTRAL UNIVERSITY LIBRARY

Stability of the column at extremes of pH must be considered when choosing binding and elution buffers. The working stability of the Resource Q column is pH 2 to 12, and the stability for cleaning is pH 1 to 14.

The flow rate for the test should be appropriate for the column to be connected. For the Resource Q column, the recommended flow rate is 1 to 10 ml/min and the maximum operating pressure is 4 bar.

3. Test for system leaks and ascertain monitor stability according to system documentation. Test accuracy of gradient composition using conductivity meter.

Adding 0.1% (v/v) acetone to the elution buffer used in step 2 will allow observation of gradient composition at 280 nm using the UV monitor. Do not include acetone in the elution buffer used for the chromatographic run.

Prepare column

4. Purge system with binding buffer to remove any air, then install column in system.

New columns may require precycling prior to use; consult product documentation. Used columns may require cleaning or removal of storage solution prior to use (see Support Protocol 4).

5. Wash column with 5 V_c (5 ml for Resource Q column) of elution buffer at a flow rate of 5 ml/min and check for leakage.

V_c refers to the packed bed volume of a column and is calculated according to the equation $V_c = \pi r^2 \times L$, where r is the radius of the column and L is the bed height (i.e., the height of the packed medium in the column). As the Resource Q column used here has a packed bed volume of 1 ml, the column is washed with 5 ml of elution buffer. Consult product literature to find V_c for other columns.

6. Equilibrate column with 5 to 10 V_c of binding buffer at 5 ml/min. Collect one fraction at the end of the equilibration stage and measure pH and conductivity to ensure the pH and salt concentration of the fraction are the same as that of binding buffer.

If the pH and salt concentration of the fraction are not the same as that of the binding buffer, continue passing binding buffer through the column until equilibration is complete.

Purify protein sample and regenerate column

7. Adjust pH and salt concentration of sample to binding conditions determined in Support Protocol 1. Filter sample prior to injection using a 0.22-μm filter to remove particulates that may clog system or column.

For initial experiments, a sample containing 25 mg total protein is appropriate.

For columns packed with beads of average diameter ≥90 μm, a 1-μm filter may be used; for columns packed with beads of average diameter 34 to 90 μm, an 0.45-μm filter may be used. The 0.22-μm filter is appropriate for columns packed with beads of average diameter <34 μm (e.g., the Resource Q). Alternatively, the sample may be centrifuged 5 min at 20,000 × g to remove particulates.

8. Inject sample into column by opening sample injection valve. Begin fraction collection.

1-ml fractions are generally sufficient.

The flow rate during sample injection should be reduced for concentrated samples. It may be increased for dilute samples.

9. After all sample is injected, wash column with 3 to 5 V_c of binding buffer at 5 ml/min to remove unbound or weakly bound materials.

The monitor signal should be allowed to approach the baseline before continuing.

BCU LASS CENTRAL UNIVERSITY LIBRARY

10. Elute with a linear gradient of 0% to 100% elution buffer in 20 V_c (20 ml).

    Sample-injection valves should be closed to reduce the system dead volume before beginning gradient.

11. Regenerate column by washing with 5 V_c of elution buffer.

12. Reequilibrate column by washing with 5 to 10 V_c of binding buffer.

13. Assay fractions for target protein. Pool fractions containing target protein for further processing.

SUPPORT PROTOCOL 1

TEST TUBE PILOT EXPERIMENT TO DETERMINE STARTING CONDITIONS FOR ION-EXCHANGE CHROMATOGRAPHY

Before beginning an ion-exchange separation (Basic Protocol 1 or 2), the pH and salt concentration required for binding the target protein and the change in pH or salt concentration required for elution must be determined empirically. Once the initial binding conditions have been determined, the available protein capacity (for batch chromatography) is determined to predict the amount of medium that must be used for a given amount of sample. Dynamic capacity (for column chromatography) may be determined as in Support Protocol 2. This protocol is most appropriate when the sample is not strictly limited—e.g., when milligram quantities of a recombinant protein are available. If sample is scarce, a column method should be substituted.

As illustrated in Figure 8.2.3, aliquots of the protein-containing sample to be purified are mixed in test tubes with small quantities of ion-exchange gel that have been equilibrated with a series of buffers of different pH. The optimal binding pH is chosen on the basis of disappearance of target protein in the supernatant at a certain point in the series, then the process is repeated with another series of test tubes containing gel equilibrated with a series of salt solutions of different concentrations. The optimal salt concentrations for binding and elution are chosen respectively, on the basis of disappearance, then reappearance at another salt concentration, of protein in the supernatant. Finally, aliquots containing increasing protein concentrations are mixed with the gel in a buffer under the optimal binding pH and salt concentrations previously determined. The available capacity of the gel is then determined on the basis of the point where unbound protein appears in the supernatant.

The stability of the target protein at extremes of pH must be considered—i.e., the pH range chosen for the purification should be compatible with retention of structure and activity. In general, this can be evaluated with a crude preparation by incubating aliquots at a series of pH values, then performing the structural or functional assay at the optimal pH value.

NOTE: The method presented here is for anion exchange with increasing salt concentration for elution. Differences for cation exchange are noted where appropriate. Select cation-exchange buffers from Table 8.2.1.

Materials

20 mM piperazine, pH 5.0, 5.5, and 6.0 (prepare from 100 mM stock)
20 mM 1,3-*bis*[tris{hydroxymethyl}methylamino]propane (bis-Tris propane), pH 6.5 and 7.0 (prepare from 100 mM stock)
20 mM Tris·Cl, pH 7.5, 8.0, and 8.5 (*APPENDIX 2E*; prepare from 100 mM stock)
Q Sepharose Fast Flow (50% slurry in 20% ethanol; Pharmacia Biotech) or equivalent anion exchange resin in appropriate buffer
Protein sample to be purified, containing known quantity of target and total protein
4 M NaCl (*APPENDIX 2E*)

RCU IASI CENTRAL UNIVERSITY LIBRARY

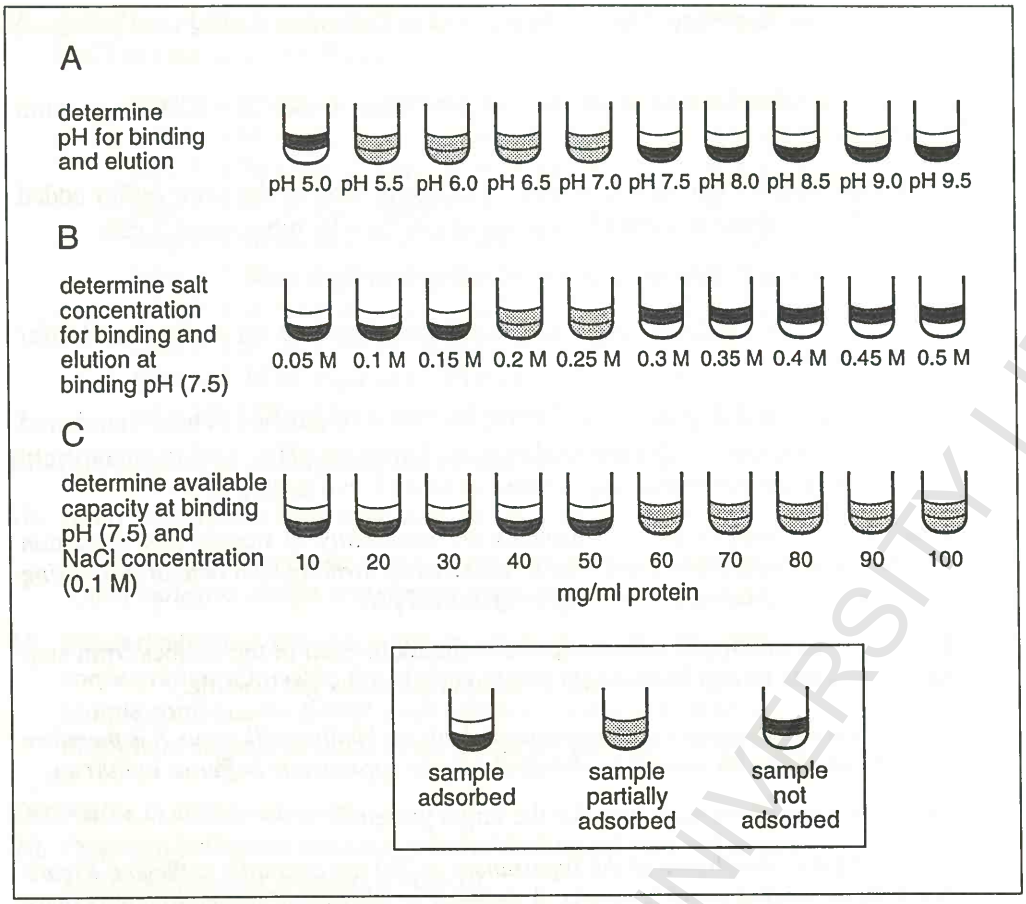

A
determine pH for binding and elution
pH 5.0 pH 5.5 pH 6.0 pH 6.5 pH 7.0 pH 7.5 pH 8.0 pH 8.5 pH 9.0 pH 9.5
B
determine salt concentration for binding and elution at binding pH (7.5)
0.05 M 0.1 M 0.15 M 0.2 M 0.25 M 0.3 M 0.35 M 0.4 M 0.45 M 0.5 M
C
determine available capacity at binding pH (7.5) and NaCl concentration (0.1 M)
10 20 30 40 50 60 70 80 90 100
mg/ml protein
sample adsorbed
sample partially adsorbed
sample not adsorbed

Figure 8.2.3 Results of a typical test tube pilot experiment for selecting initial conditions for anion-exchange chromatography. Series of tubes with protein and ion-exchange medium at (**A**) varying pH (to select binding and elution pH); (**B**) constant pH, varying NaCl concentration (to select binding and elution salt concentrations); (**C**) constant pH and NaCl concentration, varying protein concentration (to determine available capacity). The binding pH here will be ≥7.5; the elution pH will be ≤5.0; the binding salt concentration will be ≥0.15 M; the elution salt concentration will be >0.3 M; and the available capacity will be >50 mg/ml.

15-ml test tubes
Centrifuge with rotor accommodating 15-ml test tubes (optional)

Determine optimal binding pH

1. Set up a series of test tubes, numbered 1 to 8, containing different buffers as follows:

   tube 1: 5 ml 20 mM piperazine, pH 5.0
   tube 2: 5 ml 20 mM piperazine, pH 5.5
   tube 3: 5 ml 20 mM piperazine, pH 6.0
   tube 4: 5 ml 20 mM bis-Tris propane, pH, 6.5
   tube 5: 5 ml 20 mM bis-Tris propane, pH 7.0
   tube 6: 5 ml 20 mM Tris·Cl, pH 7.5
   tube 7: 5 ml 20 mM Tris·Cl, pH 8.0
   tube 8: 5 ml 20 mM Tris·Cl, pH 8.5.

2. Shake bottle of Q Sepharose Fast Flow to resuspend gel. Pour 25 ml of gel slurry into a graduated cylinder, then allow gel to settle. Adjust volume of liquid above gel to equal the volume of settled gel (i.e., prepare 50% slurry), then resuspend with a glass rod.

3. Add 2 ml of 50% slurry of Q Sepharose Fast Flow to each tube and mix.

BOU INSTITUTE CENTRAL UNIVERSITY LIBRARY

SP Sepharose Fast Flow (Pharmacia Biotech) or equivalent may be used for cation exchange.

4. Allow gel to settle to bottom of tube or centrifuge 1 min at ~5000 × *g*, room temperature.
5. Decant supernatant from each tube and replace with 5 ml of the same buffer added in step 1. Gently shake or vortex to resuspend gel, then let tubes stand 2 min.
6. Repeat steps 4 and 5 three times to equilibrate gels in each tube.
7. Decant supernatants and resuspend each equilibrated gel in 1 ml of the same buffer added in step 1.
8. Set up eight identical aliquots (≥1 ml) of the protein to be purified in tubes numbered 1 to 8, each containing 0.1 to 1 mg total protein. Adjust the pH of each to obtain eight different pH values corresponding to those of tubes 1 to 8 in step 1.

   Quantity of protein to be used depends on availability of sample and extinction coefficient of target protein. pH may be adjusted by titrating with acid or base using pH paper or by diluting 1:1 with buffer of desired pH.

9. Add the corresponding pH-adjusted protein aliquot to each of the 8 tubes from step 7. Mix gently for 10 min by periodic swirling, then allow gel to settle.

   If the salt concentration of the sample is too high, no binding will occur. It is therefore recommended that the sample be desalted into the appropriate buffer as in UNIT 8.3.

10. Assay supernatant from each tube for the target protein.

    Measuring the absorbance of the supernatant at 280 nm is usually sufficient. Figure 8.2.3 shows a set of possible results. A decrease or absence of target protein activity (or absorbance) in the supernatant indicates binding. Specific assays for the target protein may also be used.

11. Select the lowest pH at which the highest quantity of target protein is bound to use as the binding pH for further work. Select the highest pH where no binding occurs as the elution pH.

    For anion exchange, it is best to select a binding pH no more than 0.5 to 1 pH unit (pH 7.5 to 8.0 in Fig. 8.2.3) above the highest pH at which no binding occurs. If too high a pH is chosen, elution becomes difficult and high salt concentrations may be required.

    For cation exchange, select a binding pH 0.5 to 1 unit below the lowest pH where no binding occurs. Select the lowest pH where no binding occurs as the elution pH.

Determine salt concentration for binding and elution

12. Set up a series of test tubes, numbered 1 to 10. Using the binding buffer selected in step 10 in place of the different buffers in step 1, perform steps 1 to 7 to equilibrate the Q Sepharose Fast Flow with buffer in all ten tubes.
13. To each tube, add the same amount of the protein to be purified that was added to the tubes in step 8, equilibrated in the binding buffer selected in step 10.
14. Mix contents of tubes gently for 10 min, then allow gel to settle.
15. Decant supernatants from each tube and discard. Wash gel two times with 5 ml of the binding buffer, allowing the gel to settle and decanting the supernatants after each addition of buffer.

BCU IASI CENTRAL UNIVERSITY LIBRARY

16. Add 2 ml of the binding buffer to each tube, then sequentially add water and 4 M NaCl to each tube as follows:

tube 1: 1.90 ml H_2O and 0.10 ml 4 M NaCl (0.10 M NaCl final)
tube 2: 1.80 ml H_2O and 0.20 ml 4 M NaCl (0.20 M NaCl final)
tube 3: 1.70 ml H_2O and 0.30 ml 4 M NaCl (0.30 M NaCl final)
tube 4: 1.60 ml H_2O and 0.40 ml 4 M NaCl (0.40 M NaCl final)
tube 5: 1.50 ml H_2O and 0.50 ml 4 M NaCl (0.50 M NaCl final)
tube 6: 1.40 ml H_2O and 0.60 ml 4 M NaCl (0.60 M NaCl final)
tube 7: 1.30 ml H_2O and 0.70 ml 4 M NaCl (0.70 M NaCl final)
tube 8: 1.20 ml H_2O and 0.80 ml 4 M NaCl (0.80 M NaCl final)
tube 9: 1.10 ml H_2O and 0.90 ml 4 M NaCl (0.90 M NaCl final)
tube 10: 1.00 ml H_2O and 1.00 ml 4 M NaCl (1.00 M NaCl final).

17. Mix contents of tubes gently for 10 min, then allow gel to settle.

18. Assay supernatant from each tube for the target protein.

Figure 8.2.3 shows a set of possible results. Presence of target protein in the supernatant indicates elution of the target protein.

19. Select the highest salt concentration at which no elution occurs as the maximum salt concentration allowable for binding target protein and for washing away unbound sample components. Select a salt concentration at least 0.05 M above the concentration where no protein binds to the gel as the salt concentration for elution.

Determine available capacity

20. Prepare a buffer/salt solution with the binding pH selected in step 10 and the binding salt concentration selected in step 19.

21. Set up a series of test tubes numbered 1 to 10. Using the buffer/salt solution prepared in step 20 in place of the different buffers in step 1, perform steps 1 to 7 to equilibrate the Q Sepharose Fast Flow with buffer/salt solution in all ten tubes.

22. Add aliquots of sample to each tube to obtain the following amount of target protein:

tube 1: 10 mg
tube 2: 20 mg
tube 3: 30 mg
tube 4: 40 mg
tube 5: 50 mg
tube 6: 60 mg
tube 7: 70 mg
tube 8: 80 mg
tube 9: 90 mg
tube 10: 100 mg.

23. Mix gently for 10 min, then allow gel to settle.

24. Assay supernatant from each tube for the target protein.

Figure 8.2.3 shows a set of possible results. Presence of target protein in supernatant indicates that the available capacity has been exceeded.

25. Select the highest protein concentration at which the target protein does not appear in the supernatant as the available capacity.

Calculate the amount of ion-exchange medium needed for batch purification based on 50% of the available capacity (UNIT 8.1). If a column system is used in pilot experiments, however, it is best to determine the dynamic capacity of the medium (see Support

BCU ASTON CENTRAL UNIVERSITY LIBRARY

Protocol 2). However, 20% of the available capacity is a safe and reasonable starting point for pilot experiments using column chromatography.

SUPPORT PROTOCOL 2

MEASUREMENT OF DYNAMIC (COLUMN) CAPACITY AND BREAKTHROUGH CAPACITY OF ION-EXCHANGE COLUMNS

For chromatographic separations that will be performed on a routine basis as well as separations that will be scaled up, it is necessary first to optimize conditions with respect to resolution and to determine the capacity of the medium with respect to the target protein (see Support Protocol 1). Knowledge of the capacity allows optimal use of the gel medium in terms of cost and yield, and also allows working limits to be set on the sample load to ensure robustness of a particular purification step. After capacity has been determined, the separation can be optimized with respect to time. There are two capacity terms that commonly appear in the literature (in addition to available capacity; also see Support Protocol 1 and *UNIT 8.1*): dynamic capacity and breakthrough capacity.

The dynamic (or column) capacity is the capacity of an ion-exchange column under defined conditions of flow rate, pH, salt concentration and sample concentration. The breakthrough capacity (Q_B) for a system is obtained by calculating the amount of protein that has been absorbed by the column when the sample is first detected in the effluent, or when the recorder signal reaches some arbitrarily defined percent of full-scale deflection. Breakthrough capacity values at 50% full-scale deflection (Q_{B50}) are commonly reported in the literature. However, it is generally recommended to use only 10% to 20% of the published Q_B as the practical capacity of the column to ensure high recovery of the target protein.

***Additional Materials** (also see Basic Protocol 2)*

Ion-exchange gel of unknown capacity, and column
Elution buffer capable of eluting target protein in single step of salt concentration or pH (e.g., 2 M NaCl; see Support Protocol 1)

1. Pack a defined amount of the ion-exchange gel into the column.

   See UNIT 8.3 for guidelines on column packing. Normally, a quantity of gel sufficient to give a packed bed volume of ~1 ml is sufficient. In the case of a prepacked column, the amount of gel is predetermined. The packed bed volume (V_c) is calculated according to the equation $V_c = \pi r^2 \times L$, where r is the radius of the column and L is the bed height (i.e., the height of the packed medium in the column).

2. Prepare chromatography system and column (see Basic Protocol 2, steps 1 to 6) using the optimal binding and elution conditions determined in Support Protocol 1.

   Do not exceed 75% of the flow rate used for packing. The effect of flow rate on capacity should be determined in a true optimization.

3. Prepare protein solution and inject into column (see Basic Protocol 2, steps 7 and 8). Apply solution continually until the recorder shows >50% full-scale deflection, then stop.

   The absorbance of the binding buffer corresponds to 0% and the absorbance of the sample corresponds to 100% full-scale deflection.

4. Wash column with binding buffer until 0% full-scale deflection is approached.

5. Elute proteins with a buffer providing a single-step increase in salt concentration or pH. Continue collecting eluant until recorder shows ≤2% full-scale deflection, then pool fractions.

   A typical chromatogram for this procedure is presented in Figure 8.2.4.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

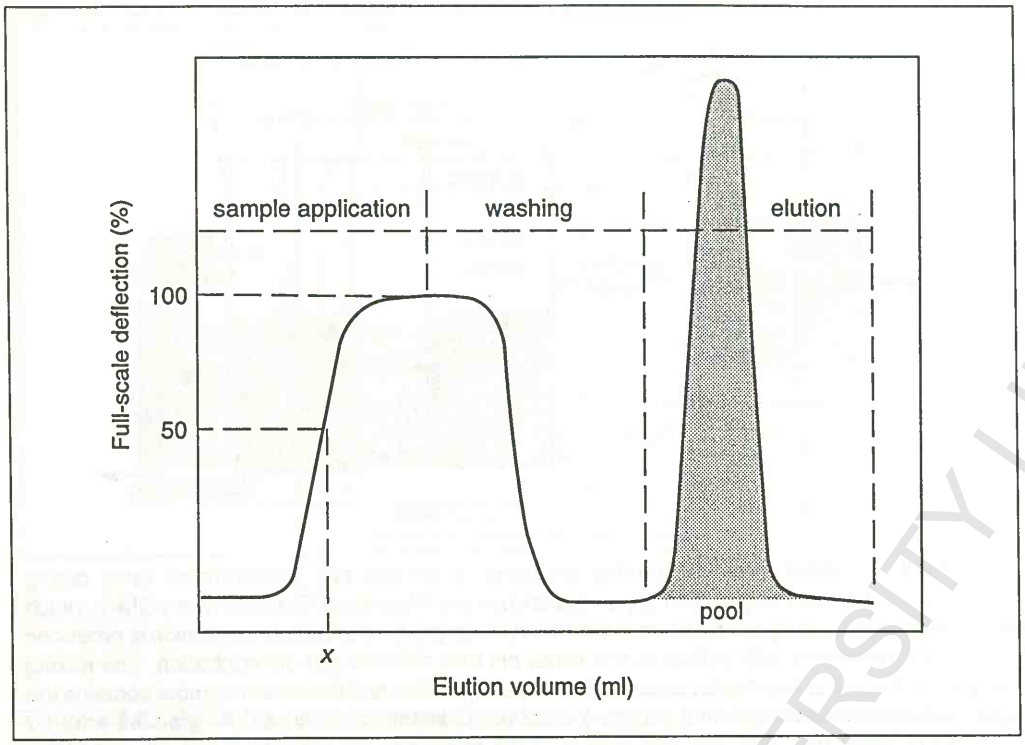


Figure 8.2.4 Typical chromatogram for determining the capacity of an ion-exchange column. The volume of eluant (x) at 50% full-scale deflection is used to calculate the breakthrough capacity (Q_{B50}).

6. Assay pooled fractions from step 5 for target protein. Calculate maximum amount of protein (A) that can be bound to the column according to the equation $A = C_p \times V$, where C_p is the protein concentration of the pooled fractions (mg/ml) and V is the volume of the pooled fractions (ml).
7. Calculate the column (dynamic) capacity (mg/ml) using the equation *dynamic capacity* $= A/V_c$.
8. Calculate breakthrough capacity (Q_{B50}) using the equation $Q_{B50} = (C_s \times x)/V_c$, where C_s is the concentration of protein in the original sample and x is the volume of sample that has been applied to the column at the point where 50% full-scale deflection has been attained.

If Q_{B50} greatly exceeds the dynamic capacity, then the target material may still be adsorbed on the column, and the elution conditions should be reexamined.

See Figure 8.2.4 for graphical depiction of these variables.

GRADIENT-FORMATION TECHNIQUES

SUPPORT PROTOCOL 3

Simple Gradient Mixers

Gradient mixers of the type shown in Figure 8.2.5 can generate linear pH or salt gradients in gravity-based or one-pump chromatography systems. Similarly designed gradient mixers used for casting electrophoresis gels may also be used for chromatography. The reservoir closest to the column is filled with the binding buffer and the second reservoir is filled with an equal volume of the elution buffer. The gradient begins when the valve between the two reservoirs is opened. Continuous mixing is required in the reservoir closest to the column to ensure gradient linearity and reproducibility.

BOWLES / GENERAL UNIVERSITY LIBRARY

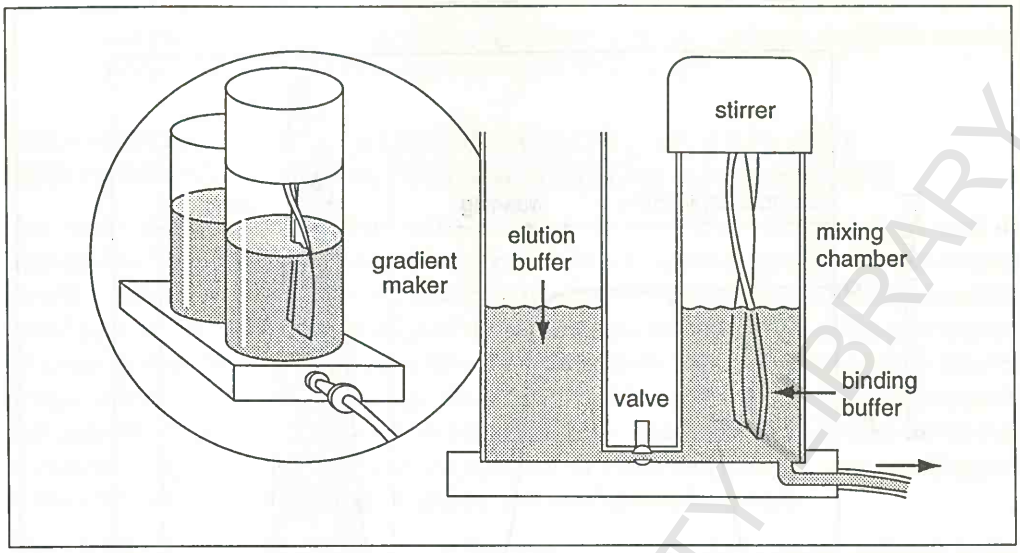


Figure 8.2.5 Gradient mixer for forming gradients of pH and salt concentration used during ion-exchange chromatography. The apparatus shown is a Pharmacia Gradient Mixer GM-1, which can be used for preparing gradients up to 500 ml. A linear gradient of salt concentration is produced by filling the reservoirs with buffers at the same pH and different salt concentration. The mixing chamber contains the lower-salt-concentration (binding) buffer, and the other chamber contains the higher-salt-concentration (elution) buffer. A continuous linear or nonlinear pH gradient may be produced by filling the reservoirs with buffers at the same salt concentration and different pH. A gradient of increasing pH is used for cation-exchange separations, and a gradient of decreasing pH is used for anion-exchange separations. The buffer chambers are joined by a channel controlled by a valve, and outflow from the mixing chamber is controlled by another valve.

Gradient Mixing with Multichannel Peristaltic Pumps

Peristaltic pumps that can accommodate three or more pumping channels simultaneously (Fig. 8.2.6) can be used to form linear or complex gradients (see Critical Parameters). Linear gradients are easily formed by using the same size tubing for all channels, pumping with one channel from the elution buffer reservoir to the binding buffer reservoir, and pumping with the other two channels from the binding buffer reservoir to the column. Continuous mixing is required in the binding buffer reservoir closest to the column to ensure gradient linearity and reproducibility.

Switch-Valve-Based Gradient Mixing

Switch-valve-based gradient formation is useful for forming step, linear, and complex gradients (see Critical Parameters). Gradients are formed by proportioning variable amounts of binding buffer and elution buffer through a single three-way valve using a system controller. When programmed to deliver 10% elution buffer, the controller will open the valve port for binding buffer 90% of the time and the port for elution buffer 10% of the time. As higher switching rates are required for lower flow rates to ensure gradient accuracy, high-quality system controllers will vary the rate at which the valve switches (i.e., the number of switching events per unit time) as a function of flow rate. A dynamic mixer should be included in the flow path just after the proportioning valve to ensure gradient accuracy. Switch-valve-based gradients are generally not very accurate from 1% to 10% elution buffer (90% to 99% binding buffer) or from 1% to 10% binding buffer (90% to 99% elution buffer). Therefore, the ionic strength of the binding and elution buffers may be adjusted to allow a gradient of the desired composition to be formed by mixing the two buffers in the range of 20% to 80% elution buffer (80% to 20% binding buffer).

RCU IASI - CENTRAL UNIVERSITY LIBRARY

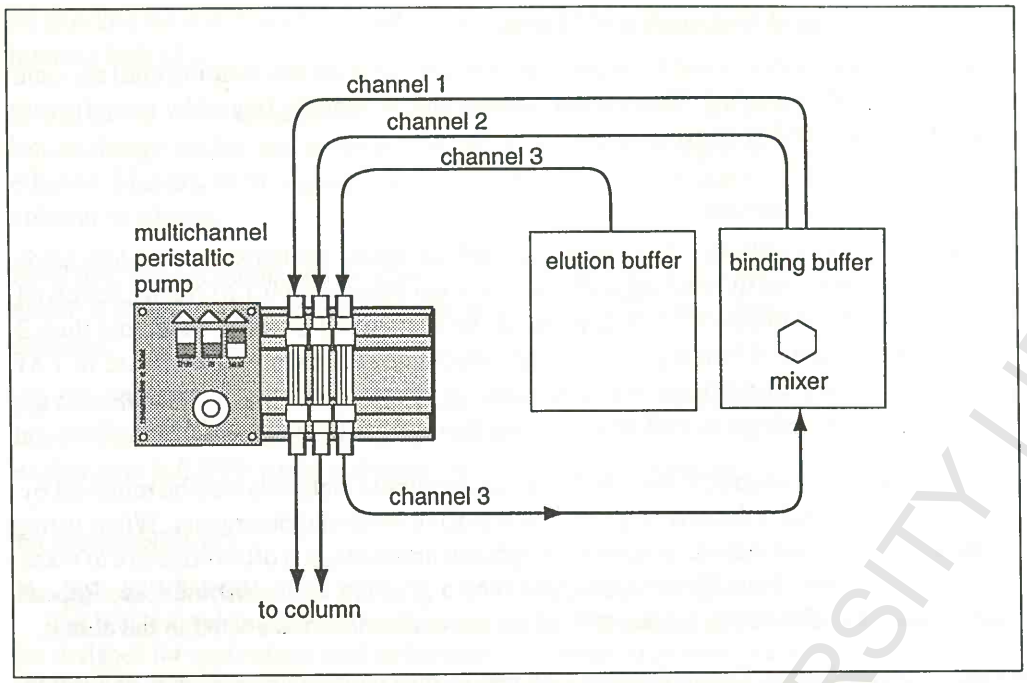


Figure 8.2.6 Peristaltic pump accommodating three pumping channels for continuous gradient formation.

Gradient Formation by Autoblending

Some chromatography systems allow gradient formation over a range of both pH and salt concentration. Four solvents—a high-pH buffer, a low-pH buffer, a concentrated salt solution, and water—are blended via four two-way valves. A system controller proportions the high- and low-pH buffers to achieve the desired pH gradient while simultaneously proportioning the concentrated salt solution and water to achieve the desired salt-concentration gradient. Gradients formed in this way are usually difficult to reproduce but may be useful for scouting initial binding and elution conditions.

Gradient Formation with Twin Pumps

Use of two positive-displacement pumps, as in HPLC and FPLC systems, provides the highest degree of accuracy, precision, and reproducibility in gradient formation. One pump delivers binding buffer while the other delivers elution buffer. A system controller allows gradient formation simply by controlling the flow rate of each pump. A dynamic mixer should be included in the flow path, after the pumps and before the column, to ensure gradient accuracy.

CLEANING AND REGENERATION OF ION-EXCHANGE MEDIA

SUPPORT PROTOCOL 4

Proper maintenance of chromatography media is as much an art as protein purification itself. Knowing what kind of contaminating material from the sample is bound to the column helps in selecting a regime for maintaining good column hygiene. It is best to practice preventative maintenance by regenerating the medium after every run, using a high-salt buffer or large change in pH, and periodically cleaning the column and checking the system flow characteristics. If any increase in column back-pressure is noted, the column should be cleaned as soon as possible. Once enough contaminant has bound to a column to reduce or block the flow, it is probably too late to save the medium without investing considerable time and effort.

BOSTON / CONTRA COSTA COUNTY LIBRARY

Routine Washing of Columns and Media

Columns or bulk media should be washed after every run with salt solution until an ionic strength of ≥2 M is reached. This should remove any substances bound by ionic forces. This is recommended after every run.

Removal of Contaminants

Alkali-soluble contaminants. Contaminants such as lipids, proteins, and nucleic acids can usually be removed by washing with 2 to 3 column volumes of 0.1 M NaOH, followed by 2 to 3 column volumes of Milli-Q purified (or equivalent) distilled water and then 2 to 3 column volumes of binding buffer. Many media can tolerate brief exposure to 1 M NaOH if harsher conditions are required to remove contaminants; however, manufacturer's guidelines should be consulted regarding pH stability of the medium.

Hydrophobic contaminants. Lipids and other hydrophobic materials may be removed by washing with alcohol solutions (e.g., 70% ethanol) or nonionic detergents. When using alcohols or organic solvents to remove hydrophobic materials, it is often effective to wash first with a gradient from 0% to 100%, and then a gradient from 100% to 0%. Repeat washing with the alternating gradients until no contaminants are detected in the eluant.

Metallic contaminants. Contamination with metals may cause blue or gray discoloration at the top of an ion-exchange column. These metals can originate from impurities in buffer salts or water, or they can leach from metallic system components in contact with buffer solutions. Metal contaminants may usually be removed by treating the column with several column volumes of 10 mM HCl (i.e., pH 2) saturated with EDTA. Manufacturer's guidelines should be checked for pH stability of medium and column before exposure to mild or strong acid solutions.

Precipitated materials. Precipitated materials that have accumulated in a column are very difficult to remove. Unless the precipitated material can be physically removed by removing the gel at the top of the column, the material will have to be resolubilized. Detergents, urea, and guanidine·Cl can be introduced to help dissolve contaminants. The viscosity of concentrated solutions used for cleaning may require very low flow rates. The column may be equilibrated and incubated in 6 M urea, then washed with distilled water and buffer. Degradative enzymes (e.g., proteases and lipases) may also be introduced in an attempt to degrade precipitates on the column; however, enzymatic cleaning can be quite expensive. It may be useful to reverse the direction of flow through the column (i.e., from bottom to top) when attempting to remove particulate material from the top of the bed.

SUPPORT PROTOCOL 5

STORAGE OF ION-EXCHANGE MEDIA

Bacterial and microbial growth can seriously interfere with the chromatographic properties of any chromatography medium and endanger the sample as well. Prior to storage, all media should be cleaned and sanitized. During prolonged experiments or storage (i.e., >24 hr at room temperature or 48 hr at 4°C), a bacteriostatic or antimicrobial agent should be added to the ion exchanger. Antimicrobials chosen to be added for storage should not bind to the ion exchanger and should be easily removed when the gel is to be reused.

Treatment with 1 M NaOH for 1 hr will sanitize most media by lowering the bacterial count 100-fold, and will help solubilize dead cells (Pharmacia Biotech, 1993). Storage of sanitized media in 0.1 M NaOH will not leave any toxic materials on the column—a significant consideration in pharmaceutical production. Manufacturer's guidelines should

BCU IAS CENTRAL UNIVERSITY LIBRARY

be checked for stability of the medium and column in NaOH. Silica-based media will not tolerate high pH.

Exposure to 70% ethanol for 3 to 4 hr can be used to sanitize most media. Most ion-exchange media and columns can be stored for prolonged periods of time in 20% ethanol. Manufacturer's guidelines should be checked for stability of the medium and column in ethanol.

Effective antimicrobial agents appropriate for use with anion-exchange media include either 0.001% phenyl mercuric salts or 0.002% chlorohexidine, in weakly alkaline solution.

An effective antimicrobial agent for use with cation-exchange media is 0.005% merthiolate in weakly acidic solution. An agent that is effective for use with either anion or cation exchangers is 0.05% trichlorobutanol in weakly acidic solution.

COMMENTARY

Background Information

The first ion exchangers were synthetic resins designed for applications such as demineralization, water treatment, and recovery of ions from waste. Early ion-exchange resins were tightly cross-linked hydrophobic polymers, highly substituted with ionic groups, and had very high capacity for small ions. The high degree of cross-linking provided mechanical strength, but the limited porosity restricted use of these media with large molecules. In addition, the high charge density resulted in very strong binding, and the hydrophobic matrix tended to adsorb and denature labile biological materials. The first ion exchangers designed for use with biomolecules were developed by Peterson and Sober (1956). These cellulose-matrix ion exchangers were very hydrophilic and had little tendency to denature proteins, but had low capacities and poor flow characteristics. Modern ion-exchange media have very high capacity, often >100 mg/ml. Modern particle technology has significantly increased the mechanical strength and flow characteristics of ion-exchange media so that separations that once required hours or days may now be completed in a few minutes.

Ion-exchange can easily be coupled to other methods of protein separation in a multistep purification. Samples purified by ion exchange can typically be applied to a hydrophobic-interaction chromatography (HIC) medium simply by adding NaCl (or other salts) to a concentration sufficient for binding (*UNIT 8.4*). Samples to be further purified by chromatofocusing or an additional round of ion exchange will require desalting (*UNIT 8.3*). Samples may usually be applied directly to gel-filtration columns (*UNIT 8.3*) with no intervening treatment.

Critical Parameters

The most critical aspects of any ion-exchange experiment are selection of the appropriate ion-exchange medium and of the initial conditions for binding and elution of the target protein or contaminants. Once these parameters are selected, conditions can be optimized with respect to the three primary parameters of any chromatographic separation—i.e., resolution, capacity, and speed. These three parameters are mutually exclusive. Resolution will typically decrease when the sample load is increased and when the flow rate is increased (to increase speed). It is generally recommended that resolution be optimized first, then capacity, and finally speed.

Optimization of resolution

This parameter is most easily optimized by controlling the selectivity (*UNIT 8.1*) of the medium through manipulation of binding and elution conditions. Construction of chromatographic titration curves followed by optimization of the shape of the elution gradient is a common approach to achieving maximum resolution.

Chromatographic titration curves can be constructed by performing a series of anion- and cation-exchange experiments at different pH values as in Figure 8.2.7. For each of these experiments, the elution salt concentrations (in mM NaCl) for sample components A, B, and C are plotted on the y axis and the operating pH is plotted on the x axis. It is then simple to observe the pH ranges within which optimal resolution between the various sample components is achieved. The different sample components are usually detected by UV absorption, but may also be viewed by other means (e.g., fluorescence or bioactivity). Data for chroma-

BALDWIN-WALLACE UNIVERSITY LIBRARY

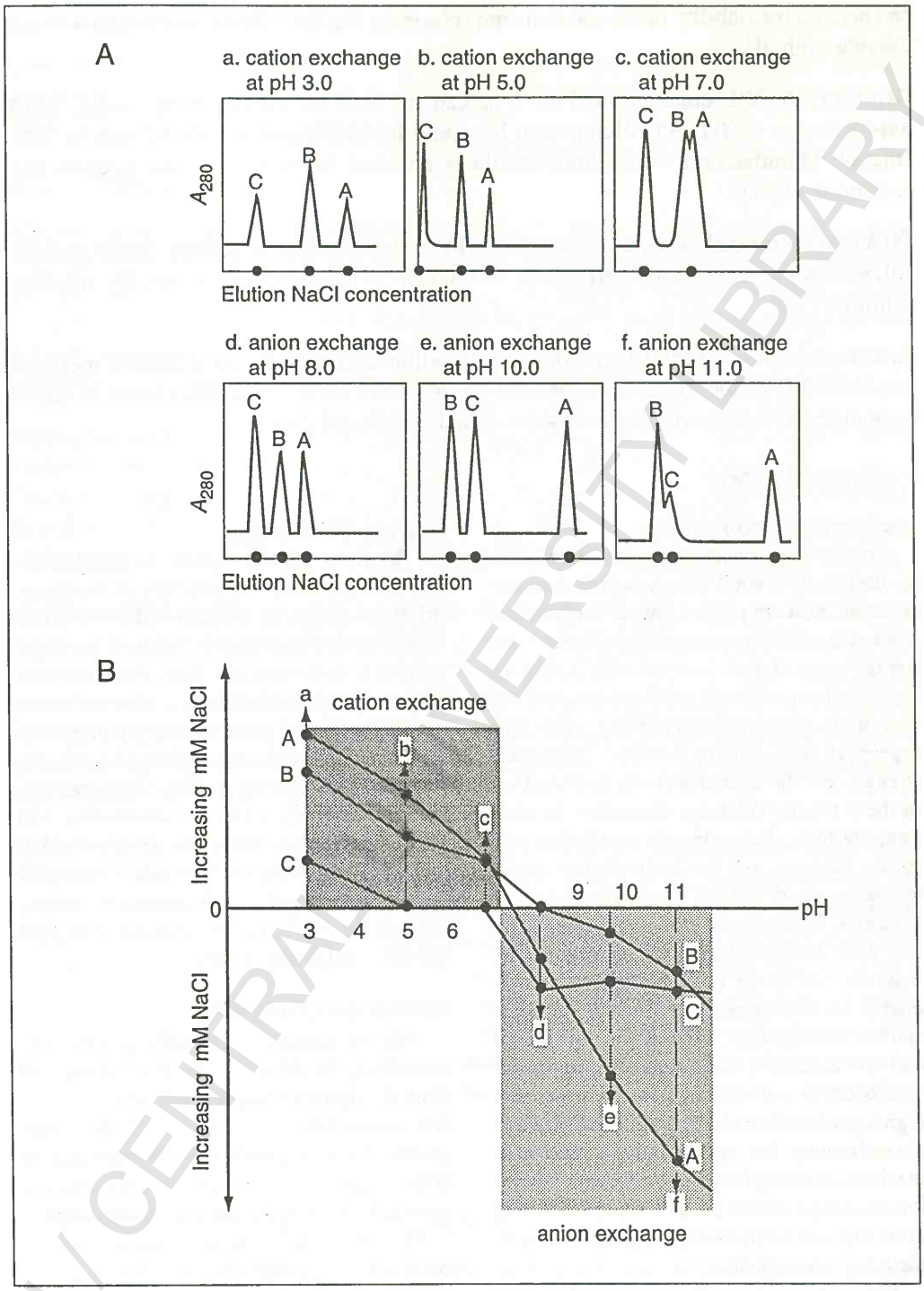


Figure 8.2.7 Construction of a chromatographic titration curve. (**A**) Six chromatographic separations (a to f) of a sample containing proteins A, B, and C are carried out at six different operating pHs using a linear gradient of increasing salt concentration. (**B**) The elution ionic strengths for components A, B, and C (*y* axis) are plotted against the operating pH (*x* axis) for experiments a to f. Note that the salt concentration in mM NaCl (*y* axis) is ascending in both directions. The optimal resolution with respect to component A occurs below pH 5 with cation exchange and above pH 9 with anion exchange.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

tographic titration curves can often be generated by fast screening with low-volume gradients (e.g., 5 to 10 V_c) and low sample mass (e.g., <1 mg).

After optimization of resolution with respect to pH, the shape of the elution gradient can be optimized. A gradient can be conceptually divided into segments, and each segment may have a different gradient shape. Gradient shapes include linear, concave, convex, and step. Complex gradients are constructed by using different gradient shapes in consecutive segments. The slope (e.g., mM NaCl per minute) of a linear segment is constant, that of a concave segment continually increases, that of a convex gradient continually decreases, and that of a step gradient approaches infinity. Resolution between sample components in any segment can usually be increased by decreasing the gradient slope in that segment. However, decreasing the gradient slope will increase the peak volume during elution, the gradient volume, and the time of the separation. Step gradients can provide the most concentrated fractions, and are desirable when minimizing sample volume is of primary importance. Use of concave gradients in the initial part of a segment and convex gradients in the later part of a segment can improve resolution. The capabilities of available equipment must be considered in the construction of complex gradients. Simple gradients that do not stress the performance limits of the system will result in robust methods that are insensitive to small changes in sample or buffer composition and gradient shape, thereby increasing probability of success.

To construct a complete gradient profile, it is first necessary to identify the gradient range where desorption of the target protein occurs, as well as the gradient slope that provides optimal resolution. A single-step elution is then done at a salt concentration just below the range of where the target protein elutes, to remove weakly bound material. The target protein is then desorbed and eluted. A single-step elution at a high salt concentration will then remove tightly bound sample components and prepare the ion-exchange gel for regeneration. This approach, using step gradients before and after the range of interest, can provide simple, robust methods that are easy to scale up, and will save time and reagents.

Different ion-exchange media offer different selectivities (*UNIT8.1*). If adequate resolution for a specific stage in a purification cannot be obtained with a specific ion-exchange medium, it may be desirable to select an alternate ion exchanger with a different functional group, or to use a different chromatographic technique at this stage in the separation. Salts other than NaCl may be used in elution, but there are no specific effects from different salts on the separation in ion-exchange chromatography, other than differences in concentration. Brewer et al. (1989) provides a good review of the effects of nonbuffering ionic species in ion exchange.

Optimization of capacity

When optimum resolution has been achieved, the effect of increasing sample load on resolution should be evaluated to determine the optimal balance of resolution and capacity utilization. The procedure for determining the dynamic (column) capacity is described in Support Protocol 3 and can be performed using the conditions for optimal resolution (as determined above) to elute the target protein. Once the capacity is determined under the optimal elution conditions, the effect of changes in sample load on resolution can be investigated to identify optimal conditions. It is generally recommended that only 10% to 20% of the dynamic capacity be used to ensure high recoveries of the target protein.

Optimization of throughput

Once optimal conditions for resolution and capacity have been determined, the flow rate in each gradient segment may be increased and the effect on resolution and capacity monitored to complete the optimization process.

Troubleshooting

Common problems encountered during ion-exchange experiments, their potential causes, and suggested solutions are outlined in Table 8.2.5.

Anticipated Results

A chromatogram illustrating typical results of an ion-exchange column chromatography experiment (i.e., of Basic Protocol 1) is shown in Figure 8.2.8.

The goal of any purification step is the pure target protein obtained at the end of the day. The chromatogram is only a record for comparison and is deceptive at best. Great care should be exercised in interpreting a chromatogram—i.e., the results from assays performed on fractions collected must be used in interpreting the chromatogram and inferring the identity of any given chromatographic peak.

PURDUE UNIVERSITY LIBRARY

Table 8.2.5 Troubleshooting Guide for Ion-Exchange Chromatography

Problem	Possible cause	Solution
The protein does not elute in the salt gradient	Buffer pH is incorrect	Use buffer pH closer to pI of the protein
	Salt concentration is too low	Use more concentrated elution buffer
Protein does not elute in pH gradient	Buffer pH is incorrect	Calibrate pH meter and prepare new solutions
Protein elutes in wash phase	Salt concentration of binding buffer is too high	Decrease salt concentration of binding buffer
	Salt concentration of sample is too high or pH is wrong	Perform buffer exchange on sample
	Column is not properly equilibrated	Repeat or prolong equilibration until conductivity is constant
	Ionic detergents or other additives are absorbed to column	Clean column
Resolution is less than expected	Gradient slope is too steep	Use shallower gradient or include plateau in gradient
	Flow rate is too high	Run separation at lower flow rate
	Proteins or lipids have precipitated on columns	Clean and regenerate column
	Sample has not been filtered properly before application to the column	Regenerate column, filter sample, and repeat chromatography
	Proteins in sample have formed aggregates that bind strongly to gel	Use urea or detergents to clean column
	Column is poorly packed	Check packing by running a colored compound and observing the band. Repack column if necessary.
	Too much sample mass has been loaded onto column	Clean and regenerate column. Decrease sample load and repeat run.
	Detector flow cell volume or mixing spaces in or after column are too large	Change flow cell or reduce all post-column volumes
Leading or very rounded peaks observed in chromatogram	Column is overloaded	Decrease sample load and repeat run
	Column is poorly packed	Check packing by running a colored compound and observing the band. Repack column if necessary.
	Column needs regeneration	Clean and regenerate column. If this does not help, replace with new one.
Tailing of peak is observed in the chromatogram	Sample is too viscous	Reduce amount of protein, or remove nucleic acids from sample.
	Precipitation of protein has occurred in column filter and/or at top of the gel bed	Clean column, exchange or clean filter.
	Column is poorly packed	Check packing by running a colored compound and observing band. Repack column if necessary.
Previous elution profile cannot be reproduced	Buffer pH and/or ionic strength are incorrect	Prepare new solutions
	Sample has deteriorated during storage	Prepare fresh sample
	Proteins or lipids have precipitated on column	Clean and regenerate column

continued

RCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 8.2.5 Troubleshooting Guide for Ion-Exchange Chromatography, continued

Problem	Possible cause	Solution
	Sample has not been filtered properly	Regenerate column, filter sample carefully and repeat step
	Equilibration was incomplete	Repeat or prolong equilibration until conductivity is constant
Recovery of activity is low but recovery of protein is normal	Target protein is not stable in the elution buffer and is therefore inactivated	Change elution conditions
	Enzyme separated from cofactor or other substance needed for activity	Test by pooling all fractions and repeating assay
	Microbial growth has occurred in the column	Clean column and store gel in 20% ethanol or other antimicrobial agent[a]
Protein amount in the eluted fractions is much less than expected	Protein was degraded by proteinases	Add proteinase inhibitors to buffers
	Protein adsorbed to filter during sample preparation	Use another type of filter or add detergents to buffer
	Microbial growth occurred in the column	Clean column and store gel in 20% ethanol or other antimicrobial agent[a]

[a]Microbial growth rarely occurs in columns during use, but steps should always be taken to prevent infection of packed columns, buffers, and gel suspensions; see Support Protocol 5.

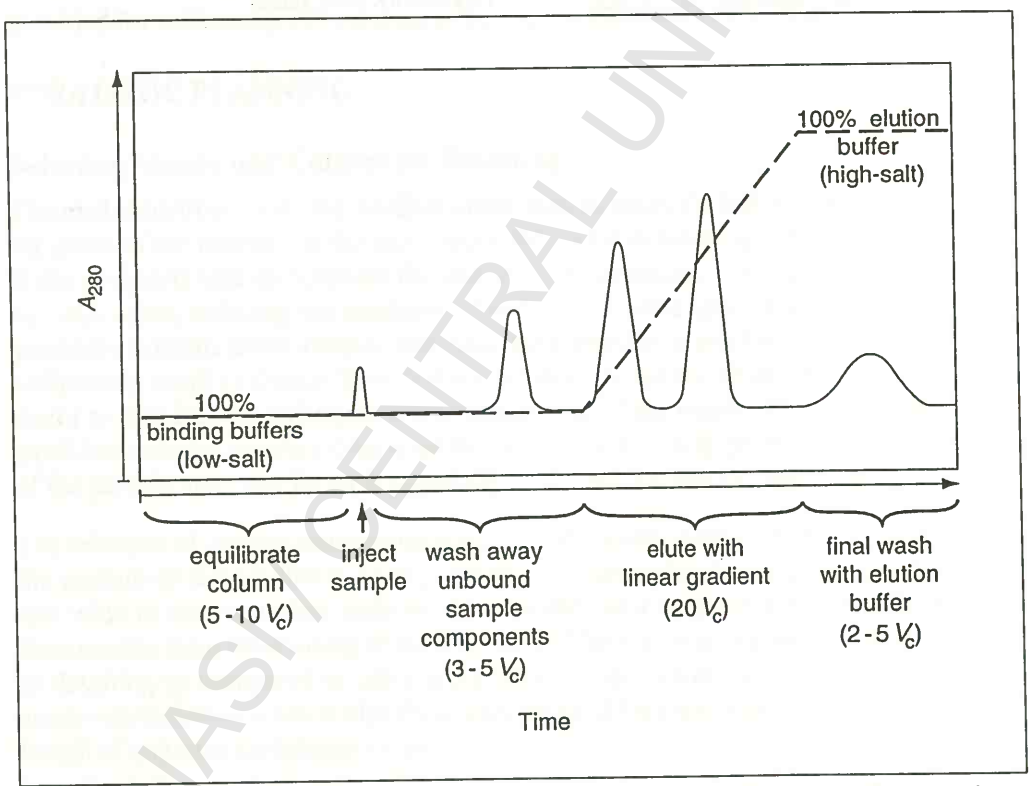

100% elution buffer (high-salt)
A_{280}
100% binding buffers (low-salt)
equilibrate column (5-10 V_c)
inject sample
wash away unbound sample components (3-5 V_c)
elute with linear gradient (20 V_c)
final wash with elution buffer (2-5 V_c)
Time

Figure 8.2.8 Chromatogram representing typical results of stages of an ion-exchange column chromatography experiment. A linear gradient of 0% to 100% elution buffer is used as in Basic Protocol 2. The broken line superimposed on the chromatogram represents the composition of the elution gradient (100% binding buffer to equilibrate column and remove unbound sample components, followed by a linear gradient from 100% binding buffer to 100% elution buffer to elute bound proteins and a final wash with 100% elution buffer). V_c represents the packed bed volume of the column.

UNIVERSITY OF CENTRAL ASIA LIBRARY

Time Considerations

If all necessary equipment is available and functioning properly, it should be possible to develop a chromatographic method in a day. With automated systems (e.g., HPLC and FPLC) and modern ion-exchange media, method development may require just a few hours. However, the most time-consuming aspect in developing a purification step is typically the determination of protein purity in the fractions collected, which will always depend on the specific protein of interest. The key to minimizing the time required for such determinations is to strive for high resolution and a minimum number of fractions.

Literature Cited

Brewer, J., Karlsson, E., and Rydén, L. 1989. Ion exchange chromatography. *In* Protein Purification: Principles, High Resolution Methods and Applications (J.C. Janson and L. Rydén, eds.) pp. 107-149. VCH Publishers, New York.

Gianazza, E. and Righetti, P.G. 1980. Size and charge distribution of macromolecules in living systems: *J. Chromatog.* 193:1-8.

Peterson, E.A. and Sober, H.A., 1956. Chromatography of proteins. I. cellulose ion exchange adsorbents. *J. Amer. Chem. Soc.* 78:751-755.

Pharmacia Biotech. 1993. Technical Note 211: Sanitization of BioProcess Glass Columns with Sodium Hydroxide. Pharmacia Biotech AB, Uppsala, Sweden.

Key References

Cooper, E.H., Turner, R., Webb, J.R., Lindblom, H., and Fägerstam, L. 1985. Fast liquid protein chromatography scale-up procedures for the preparation of low molecular weight proteins from urine. *J. Chromatogr.* 327:269-277.

Good example of methods development with respect to optimization of resolution.

Pharmacia Biotech. 1985. FPLC Ion Exchange and Chromatofocusing: Principles and Methods. Pharmacia Biotech AB, Uppsala, Sweden.

Contains detailed discussions of experimental approach, methodology, and applications for protein purification.

Pharmacia Biotech. 1991. Ion Exchange Chromatography: Principles and Methods, 3rd ed. Pharmacia Biotech AB, Uppsala, Sweden.

Concise descriptions of theory and practice in planning and implementing ion-exchange purification.

Contributed by Alan Williams
Pharmacia Biotech
Piscataway, New Jersey

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Gel-Filtration Chromatography

Gel filtration (GF) separates proteins solely on the basis of differences in molecular size. Separation is achieved using a porous matrix to which the molecules, for steric reasons, have different degrees of access (i.e., smaller molecules have greater access and larger molecules less). The matrix is packed into a chromatographic column, the sample mixture applied, and the separation accomplished by passing an aqueous buffer (i.e., the mobile phase) through the column. Protein molecules that are confined in the volume outside the matrix beads will be swept through the column by the mobile phase and the sample is thus eluted in decreasing order of size. The protein zones eluted are detected by an in-line UV monitor and fractions are collected for subsequent specific analysis or further preparation steps.

GF is often used to condition samples prior to further purification by adsorptive techniques. It is also frequently used in the final polishing stage (*UNIT 8.1*) to remove a few remaining contaminants with surface properties similar to the target protein (e.g., for polishing monomers from oligomers).

Three basic protocols are provided in this unit, corresponding to the major applications of gel filtration. GF is used for desalting or group separation (see Basic Protocol 1), in which the target protein and contaminating solutes differ substantially in molecular size. It is also used for protein fractionation (see Basic Protocol 2), where proteins with only a small size difference must be separated. A third application is in characterizing the molecular dimensions of proteins (see Basic Protocol 3). A Support Protocol is also provided for calibrating GF columns to be used in estimating molecular size.

STRATEGIC PLANNING

Selecting Matrix and Column for Desalting

The exclusion limit—i.e., the smallest-sized protein molecule that will be excluded from the pores of the matrix—is the most important factor in selecting a matrix for desalting. If the protein is able to permeate the matrix, flow-dependent zone broadening (see *UNIT 8.1*) will occur, reducing the resolution. On the other hand, gels of very small pore size generally contain more matrix, and thus have smaller separating volumes. A gel of sufficiently small exclusion limit, but not smaller, should therefore be selected. Another factor to take into consideration is the particle size of the matrix. This should not be too small because the pressure drop over the column is inversely proportional to the square of the particle size; this factor is especially important for gravity flow columns.

The selection of column dimensions (i.e., the bed volume of the column) is influenced by the amount of sample that is to be processed. Columns for desalting are typically short and wide to accommodate large sample volumes and keep the processing time short. Because the zone broadening of the column will have a limited effect on the peak width in desalting as compared to other applications of gel filtration—e.g., protein fractionation—short columns run at high flow rates are used for large-scale desalting. The typical length of columns for laboratory use is ~10 cm.

Columns for desalting vary greatly in size and sophistication, ranging from Pasteur pipets plugged with glass wool to prepacked high-performance GF columns. The scale of operation (*UNIT 8.1*) for which they are used also varies widely, with applications ranging from micropreparative cleanup of a 50-µl reduction/alkylation mixture using a 0.9-ml column to industrial de-ethanolization of 12 liters of human serum albumin using a 75-liter

Copyright © 1995 by John Wiley & Sons, Inc.

BAUHAUS CENTRAL UNIVERSITY LIBRARY

Table 8.3.1 Gel Filtration Matrices Suitable for Desalting[a]

Gel type	Exclusion limit		Supplier[d]
	Daltons[b]	radius, Å[c]	
Sephadex G-10	700	7	PB
Bio-Gel P-2	1,800	10	BR
Sephadex G-25	5,000	14	PB
Bio-Gel P6DG	6,000	15	BR
Sephadex G-50	30,000	25	PB
Bio-Gel P-30	40,000	28	BR

[a]Data as reported by manufacturers.

[b]Molecular mass of smallest protein excluded from matrix.

[c]Apparent radius of smallest spherical protein excluded from matrix, calculated as $R = 0.808 M_r^{1/3}$, where R is the radius and M_r the molecular mass (Hagel, 1989).

[d]Abbreviations: BR, Bio-Rad; PB, Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

column. Incidentally, columns for both these procedures may be packed with Sephadex G-25—of course, of different particle size.

Some gel-filtration media suitable for laboratory-scale desalting are listed in Table 8.3.1, and some empty columns that can be used with these media are listed in Table 8.3.2. A partial compilation of disposable prepacked columns for desalting is found in Table 8.3.3. This table also contains information about spin columns for use with a centrifuge. Spin columns require special techniques, and they are presently not widely used for desalting proteins; therefore no protocols for their use have been included in this unit. More information about spin columns can be obtained from the suppliers.

Selecting Matrix and Column for Protein Fractionation

Suitability of a gel for protein fractionation may be judged from "typical runs" made with synthetic protein mixtures, the results of which are published by suppliers. The solute of interest should ideally be eluted in the first part of the chromatogram, but not too close to the void volume. A matrix with a distribution coefficient (see Basic Protocol 3) of ~0.2 to 0.4 has been reported to yield maximum resolution (Hagel, 1989). This corresponds to an elution volume of ~1.5 to 2 times the void volume. A long bed and a matrix with a large pore volume will increase the peak-to-peak distance between the proteins and thus increase the resolution of the system (also see Critical Parameters and *UNIT 8.1*). Another way to increase resolution is to employ a GF matrix of small particle size, which will reduce the zone width of the protein bands. The flow rate will also affect resolution. As in the case of desalting, the maximum sample volume that can be processed depends upon the bed volume and the pore volume of the packed bed. Thus, there are many parameters to consider when optimizing protein fractionation by GF. In the worst case, a suitable matrix may have to be found by trial and error.

Commercial prepacked columns for protein fractionation often sacrifice column length and pore volume in favor of small particle size. This results in columns that are optimized for fast separation at the expense of some reduction in the maximum sample volume that can be loaded. Therefore, traditional columns packed in the laboratory are often still a good alternative for purification of large volumes of sample. If time is critical in the separation (e.g., in proteinase removal), fast column operation is of course advantageous. Matrices of large particle size will permit protein fractionation at low operating pressure.

RCU IAS/CENTRAL UNIVERSITY LIBRARY

Some matrices suitable for protein fractionation are listed in Table 8.3.4, and some empty columns suitable for use with these media are listed in Table 8.3.5. Table 8.3.6 lists some prepacked columns for protein fractionation. Columns and media are selected on the basis of the difficulty of the separation (i.e., how small the size difference is between the proteins to be separated) and whether scaleup will be required (i.e., the availability of a variety of columns with different diameters that can be packed with bulk media facilitates linear scaleup).

Table 8.3.2 Empty Columns Suitable for Desalting[a]

Column	Minimum bed volume (ml)	Maximum bed volume (ml)	Maximum operating pressure (MPa)	Supplier[b]
PD-10[c]	1	15	—	PB
Econo-Pac[c]	1	20	—	BR
C series[d]	0	520	0.1	PB
Econo-Column[d]	1	1374	0.1	BR
XK series[d]	2	1884	0.5	PB

[a]Data as reported by the manufacturers.

[b]Abbreviations: BR, Bio-Rad; PB, Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

[c]Plastic, disposable.

[d]Borosilicate glass.

Table 8.3.3 Prepacked Columns Suitable for Desalting[a]

Prepacked column	Gel type	Exclusion limit		Maximum sample volume (ml)	Supplier[d]
		Daltons[b]	radius, Å[c]		
Columns for use with a pump					
Hi Trap Desalting	Sephadex G-25 SF	5,000	14	1.5	PB
Fast Desalting HR 10/10	Sephadex G-25 SF	5,000	14	3.0	PB
Econo-Pac P6	Bio-Gel P-6	6,000	15	3.0	BR
Gravity flow columns					
NAP-5	Sephadex G-25 M	5,000	14	0.5	PB
NAP-10	Sephadex G-25 M	5,000	14	1.0	PB
PD-10	Sephadex G-25 M	5,000	14	2.5	PB
Econo-Pac 10DG	Bio-Gel P-6DG	6,000	15	3.0	BR
Spin columns					
MicroSpin S-200	Sephacryl S-200 HR	250,000	51	0.01-0.05	PB
Bio-Spin 6	Bio-Gel P-6	6,000	15	0.05-0.1	BR
Nick Spin Column	Sephadex G-25 SF	5,000	14	0.075-0.15	PB
Bio-Spin 30	Bio-Gel P-30	40,000	28	0.05-0.1	BR

[a]Data reported by the manufacturer.

[b]Molecular mass of smallest protein excluded from matrix.

[c]Apparent radius of smallest spherical protein excluded from matrix, calculated as $R = 0.808\ M_r^{1/3}$, where R is the radius and M_r is the molecular mass (Hagel, 1989).

[d]Abbreviations: BR, Bio-Rad; PB, Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

PSU HASLER CENTRAL UNIVERSITY LIBRARY

A matrix of small particle size will permit preparation of columns of high efficiency (i.e., low zone broadening) at the expense of high pressure drops and greater cost. Also, smaller particles generally have smaller pore volumes. The net effect of the relationships between these parameters is that the prepacked columns listed in Table 8.3.6 have resolving power of the same order of magnitude, although protein fractionation may be carried out faster on the columns that have particles of small size.

A final factor to consider in choosing a matrix for protein fractionation is the likelihood of matrix-solute interactions (e.g., as in the case of lectins on gels containing glucose-type structures).

Selecting Matrix and Column for Molecular Size Determination

A matrix with high resolving power is recommended for determination of molecular size as well as for protein fractionation (see Table 8.3.4, Table 8.3.5, and Table 8.3.6 for

Table 8.3.4 Gel Filtration Matrices for Protein Fractionation or Size Determination[a]

Gel type	Fractionation range (Daltons)	Particle size (μm)	Supplier[b]
Toyopearl HW 40S	100-10,000	25-40	TH
Superdex 30 prep grade	200-10,000	34	PB
Toyopearl HW 50S	500-80,000	25-40	TH
Sephacryl S-100 HR	1,000-100,000	47	PB
Toyopearl HW 55S	1,000-700,000	25-40	TH
Superdex 75 prep grade	3000-70,000	34	PB
Sephacryl S-200 HR	5,000-250,000	47	PB
Superose 6 prep grade	5,000-5,000,000	34	PB
Superdex 200 prep grade	10,000-600,000	34	PB
Sephacryl S-300 HR	10,000-1,500,000	47	PB
Sephacryl S-400 HR	20,000-8,000,000	47	PB
Sephacryl S-500 HR	20,000-30,000,000	47	PB
Toyopearl HW 65S	50,000-5,000,000	25-40	TH
Toyopearl HW 75S	500,000-50,000,000	25-40	TH

[a]Data as reported by manufacturers.

[b]Abbreviations: PB, Pharmacia Biotech; TH, Toso Haas. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

Table 8.3.5 Empty Columns Suitable for Protein Fractionation or Size Determination[a]

Columns	Dimensions[b] (length × inner diameter, cm)	Maximum operating pressure (MPa)	Supplier[c]
HR series	20 × 0.5–50 × 1.6	3–5	PB
G series	25 × 1.0–100 × 4.4	3-7	AM
XK series	20 × 1.6–100 × 5	0.5	PB

[a]Data as provided by manufacturers. All columns are made of borosilicate glass.

[b]Column dimensions given here are those suitable for protein fractionation and size determination. Other column dimensions may also be available.

[c]Abbreviations: AM, Amicon; PB, Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in the *SUPPLIER APPENDIX*.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 8.3.6 Prepacked Columns for Protein Fractionation or Size Determination[a]

Column type[b]	Fractionation range (Daltons)[c]	Particle size (μm)	Dimensions available (length × inner diameter, cm)	Supplier[d]
Superdex Peptide	100-7,000	13	30 × 1	PB
HiLoad Superdex 30 prep grade	200-10,000	34	60 × 1.6 60 × 2.6	PB
TSK SW 2000	500-60,000	10	30 × 0.75 60 × 0.75	TH
HiLoad Sephacryl S-100 HR	1,000-100,000	47	60 × 1.6 60 × 2.6	PB
TSK SW 3000	1,000-300,000	10	30 × 0.75 60 × 0.75	TH
Protein-Pak 60	2,000-30,000	10	30 × 0.78	WA
Superdex 75 HR 10/30	3,000-70,000	13	30 × 1	PB
HiLoad Superdex75 prep grade	3,000-70,000	34	60 × 1.6 60 × 2.6	PB
Bio-Sil SEC 125	5,000-100,000	10	60 × 0.75 60 × 2.15	BR
$G2000SW_{XL}$	5,000-150,000	5	30 × 0.75	TH
HiLoad Sephacryl S-200 HR	5,000-250,000	47	60 × 1.6 60 × 2.6	PB
TSK SW 4000	5,000-1,000,000	13	30 × 0.75 60 × 0.75	TH
Superose 6 HR 10/30	5,000-40,000,000	13	30 × 1	PB
Protein-Pak 125	10,000-80,000	10	30 × 0.78	WA
Glass GF-250	10,000-250,000	4	30 × 1	DU
Bio-Sil SEC 250	10,000-300,000	10	60 × 0.75 60 × 2.15	BR
$G3000SW_{XL}$	10,000-500,000	5	30 × 0.75	TH
Superdex 200 HR 10/30	10,000-600,000	13	30 × 1	PB
HiLoad Superdex 200 prep grade	10,000-600,000	34	60 × 1.6 60 × 2.6	PB
HiLoad Sephacryl S-300 HR	10,000-1,500,000	47	60 × 1.6	PB
$G4000SW_{XL}$	20,000-10,000,000	7	30 × 0.75	TH
Glass GF-450	25,000-900,000	6	30 × 1	DU

[a]Data as reported by the manufacturers.

[b]In general, matrices of natural polymers (e.g., Superdex, Sephacryl, or Superose) have larger pore volumes than silica-based materials (e.g., TSK SW, Protein-Pak, Bio-Sil SEC, or GF-250/450).

[c]Selectivity of materials (see *UNIT 8.1*) decreases with increased width of the fractionation range and also with decreased pore volume.

[d]Abbreviations: BR, Bio-Rad; DU, DuPont; PB, Pharmacia Biotech; TH, Toso Haas; WA, Waters. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

BU ASI / CENTRAL UNIVERSITY LIBRARY

matrices and columns suitable for both these applications). However, certain other criteria that are important for selecting a matrix and column for size determination differ from those for protein fractionation. First, when analyzing samples containing several proteins of varying molecular size (e.g., protein digests) a matrix of sufficient separation range is needed. As selectivity (see *UNIT 8.1*) inherently limits separation range, the separation range may not be optimal using a matrix of high selectivity. Second, because peak broadening is not critical as long as the peaks are resolved, short (e.g., 30-cm) columns operated at high flow rates may be used. However, in situations where the effluent is continuously monitored with a mass-sensitive detector (see Basic Protocol 3), complete resolution between peaks is essential. In this case the criteria used for protein fractionation may be applied.

BASIC PROTOCOL 1

DESALTING (GROUP SEPARATION)

Desalting, or group separation, is used to separate the target protein from low-molecular-mass contaminants. The purpose might be to remove salt prior to lyophilization, or to remove low-molecular-mass reagents—e.g., a reductive cleavage mixture from a preparation step (*UNIT 11.1*) or polybuffer from chromatofocusing. Desalting is also used to change from one buffer to another—e.g., to decrease ionic strength prior to ion-exchange chromatography (*UNIT 8.2*). The pore size of the matrix is selected to exclude the protein of interest while allowing maximum permeation of the contaminants. This is beneficial in three ways. First, because the protein does not penetrate into the pores, little dilution of the sample will take place except where the sample volume is very small. Second, because the protein is unable to diffuse into the pores, no flow-sensitive zone broadening (*UNIT 8.1*) that results from slow mass-transfer will take place. Hence, the flow rate may be kept high during the desalting. Third, because the protein and the low-molecular-mass solutes are maximally separated, very large sample volumes may be applied before contamination of the protein band becomes evident. Thus, very large sample volumes may be applied and desalted in a short time without dilution of the sample.

The procedure described here uses a closed column system and a pump; annotations explain variations needed in some of the steps if an open-bed column and gravity flow are used. The GF matrix is swelled, the column is packed, and the void volume and total volume of the column are determined in preliminary runs using high- and low-molecular-weight solutes. This information is used to determine the maximum sample volume to apply for the unknown protein to be desalted. Guidelines for estimating approximate sample volumes to use for desalting are also given. Alternative steps are provided for using this procedure either to determine elution volume or to desalt a protein for preparative purposes (or further analysis). Elution volumes must be determined to ascertain the exact void volume and total volume of a column in preparation for desalting. It is also necessary to measure the elution volume for certain protein analyses—e.g., for determining molecular size (also see Basic Protocol 3).

Materials

GF matrix (Table 8.3.1) or prepacked GF column (Table 8.3.3) with appropriate exclusion limit for protein of interest (also see Strategic Planning and Critical Parameters)
GF desalting buffer (see recipe)
Colored marker: 0.2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B_{12}
Void marker (see recipe)
Total liquid volume marker (see recipe)
Protein sample to be desalted
90°C water bath (optional)

RCU IAS/CENTRAL UNIVERSITY LIBRARY

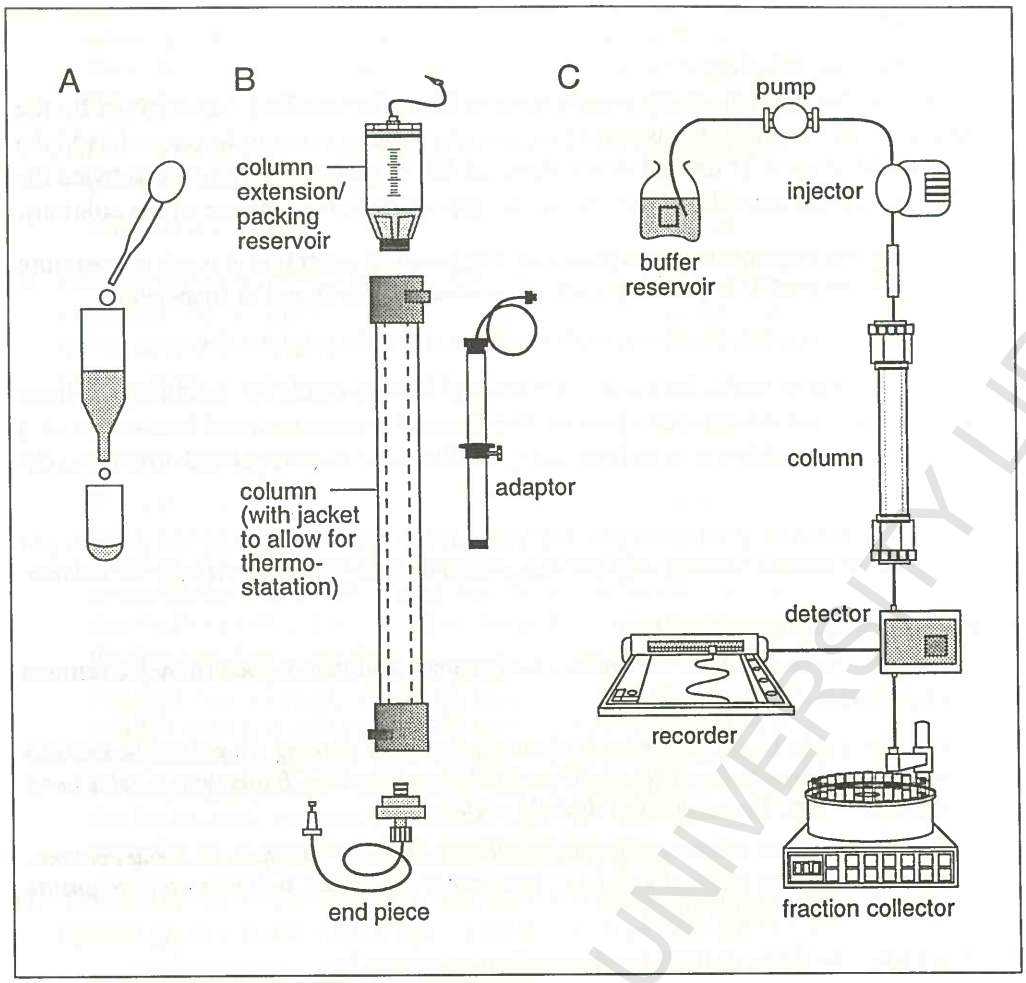


Figure 8.3.1 Equipment used for gel filtration. (**A**) Simple setup for desalting using an open-bed column made with a Pasteur pipet. (**B**) Column and attachments. (**C**) Complete automated chromatography system. Courtesy of Pharmacia Biotech.

Buchner or sintered-glass funnel
GF chromatography column (see Critical Parameters and Table 8.3.2) with column extension (optional), adaptors, and buffer reservoir (Fig. 8.3.1)
Carpenter's level
Peristaltic pump (Fig. 8.3.1)
0.22-µm filter: any 0.22-µm filter for buffer; protein-compatible for sample
Detector (optional; Fig. 8.3.1)
Chart recorder (optional; Fig. 8.3.1)
Sample applicator: sample application loop (e.g., Superloop, Pharmacia Biotech) or syringe
Fraction collector (optional; Fig. 8.3.1)

NOTE: Small-scale desalting of multiple samples may be carried out using spin columns and a centrifuge. At present, spin columns are primarily used for preparing microliter-sized samples of oligonucleotides; however, they are equally useful for small-volume protein samples. As the procedure is different from that described here, manufacturer's instructions should be consulted (see Table 8.3.3 for suppliers).

BOSTON CENTRAL UNIVERSITY LIBRARY

Prepare gel

If GF matrix is supplied dry:

1a. Calculate the amount of dry matrix needed from the swelling factor given by the manufacturer and the approximate bed volume of the column to be packed. Add the calculated amount of dry gel to a volume of GF desalting buffer equal to twice the calculated final gel volume (i.e., twice the approximate bed volume of the column).

2a. Carefully stir suspension with a glass rod. Let gel swell overnight at room temperature or for 3 hr in a 90°C water bath, stirring as needed to keep gel in suspension.

Do not use a magnetic stirrer, as this may break fragile gel particles.

3a. Allow gel bed to settle, then remove fines and broken beads by decanting the hazy solution. Repeat decantation (four to five times if necessary) until bed settles as a sharp zone, then dilute so that final slurry is 50% (v/v) settled gel and 50% (v/v) GF buffer.

Careful removal of fine particles will prevent the nets that support the column end pieces from being blocked, and will also permit a higher flow rate through the column.

If GF matrix is supplied preswollen:

1b. Wash gel with an excess of GF buffer on a Buchner or sintered-glass funnel to remove storage buffer.

In some products, the gel slurry is obtained by simply shaking the gel bottle, and this can be used to pack the column without further preparation. If this type of gel is used, skip steps 1b to 3b and proceed directly to step 4.

Gel particles may adhere to the glass wall of the funnel; therefore, to avoid contamination from other types of gels (e.g., ion-exchange media), it is important to ensure that glassware used is clean.

2b. If gel contains fines, remove by decantation as in step 3a.

3b. Dilute so that final slurry is 50% (v/v) settled gel and 50% (v/v) GF buffer.

Pack column

4. Examine column to ensure that it is clean and that the support nets are undamaged. If necessary, supply column with new support nets.

5. Mount column on a stable laboratory stand, using a carpenter's level to ensure that it is vertical. Equip column with an extension that together with the column will hold the complete volume of the gel slurry (i.e., twice the volume of the settled gel).

If <50% of the column volume is to be packed, there is no need for a column extension.

6. Remove air from outlet tubing and end piece of column by injecting GF buffer into outlet tubing (using syringe or peristaltic pump) until support net is covered with 0.5 cm of buffer. Close outlet of column.

7. Inject GF buffer into inlet tubing of adaptor until net is wetted, while holding outlet upwards to ensure that air easily escapes from net. Place adaptor in a beaker containing GF buffer until it is time to use it.

8. Swirl gel slurry from step 3a or b and pour entire slurry into column along a glass rod tilted toward the inner wall of the column or column extension. Fill the remaining space with GF buffer by carefully pouring buffer along glass rod so as to cause minimum disturbance of the gel layer. Put lid on column extension (or put top adaptor on column).

Small columns (e.g., 1 cm i.d., 5-cm length) to be used for preliminary experiments

RCU IASI CENTRAL UNIVERSITY LIBRARY

where optimum resolution is not critical may be quickly prepared by adding the gel slurry to the chromatographic tube, letting the gel settle by gravity, and washing the bed with one column volume of GF buffer before applying the sample.

9. Fill buffer reservoir with GF buffer and place at a level higher than the pump. Connect reservoir to pump using a large-i.d. tube.

   This step is to ensure a free flow of buffer to the pump head.

10. Purge pump with GF buffer and attach outlet from pump to inlet of column or column extension. Open outlet of column and start pump at flow rate recommended by manufacturer for the gel/column combination used. Continue flow until height of gel bed becomes constant (typically 1 to 4 hr).

    If no manufacturer's recommendation is available, use a flow rate 50% higher than that to be used in the separation.

    If column is to be packed by gravity flow—i.e., by attaching the buffer reservoir directly to the inlet of the column—the actual flow rate must be monitored (e.g., by continuously weighing the effluent, or in the simplest case, by using a graduated cylinder) or controlled by a pump placed after the column. In the latter case it is necessary to check that the free gravity flow exceeds the pump flow—i.e., that the pump is serving to reduce the flow to a constant value.

    The flow rate caused by gravity depends upon the hydrostatic pressure, determined by the difference in height between the air inlet in the buffer reservoir and the outlet tubing from the column. However, because the pressure drop over the column increases with the height of the packed bed, the free gravity flow rate will decrease during packing if the hydrostatic pressure is kept constant. Gravity flow rate may be increased by lowering the outlet tube or raising the buffer reservoir.

11. Turn pump off, close column outlet, remove column extension, adjust bed height if necessary, and adjust inlet adaptor so that it is flush with the bed surface.

    Avoid trapping air under the support net.

    Adjustment of the bed height is recommended only to remove excess gel. This is done by carefully swirling the upper part of the gel surface with a spatula and sucking the excess slurry out with a Pasteur pipet or tubing attached to a pump. If a crater or uneven surface is produced, the gel should be swirled so that a sharp bed surface is obtained before adjusting the inlet adaptor atop the bed. If the bed is too short the column should be repacked, although in most cases the old gel can be reused and added to new slurry prepared as in steps 1a to 3a or 1b to 3b before repacking.

12. Reattach column to pump, open column outlet, and resume flow conditions used in step 10 for 1 hr to stabilize column bed height. Readjust the adaptor so that it is flush with the new bed surface.

    In some cases it is desirable to lower the inlet adaptor a few millimeters into the gel surface to make sure that there is contact between adaptor and gel.

    Stabilizing the bed at a higher flow rate than that to be used during the separation will prevent the bed from shrinking as a result of "after-packing" during the run. If a gap appears between the adaptor and the bed surface, step 12 must be repeated at an increased flow rate before experiments are carried out. A space between the adaptor and the bed surface will act as a mixing chamber; the effect on separation will depend on the separation problem at hand.

13. Inspect packed bed visually for cracks, trapped air, and particle aggregates by observing the light scattered from a light source held behind the column. Chromatograph a colored marker (2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B_{12}) as in steps 20 to 23, and ascertain that the zone produced is horizontal and sharp.

    Quality of the packed bed need not be as high for desalting as for protein fractionation

B.C.U. „M. EMINESCU" IAŞI
CHITIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

(see Basic Protocol 2); therefore evaluation of the packed bed may be performed by visual inspection. However, it is important for it to be free of cracks and for a colored sample to produce a sharp horizontal zone.

Air bubbles would be a major problem and should be avoided at all costs (see Troubleshooting).

If column is to be stored, GF buffer containing 0.02% (w/w) sodium azide or 20% (v/v) ethanol as antibacterial agents should be prepared, then two column volumes of this mixture should be pumped through the column. The outlet is then closed and the column stored in the buffer/antibacterial agent at room temperature for several months. See recipe for GF buffer for precautions concerning azide.

Prepare and test the GF system

14. Calculate amount of GF buffer necessary for the run and filter this amount plus a 50% excess through a 0.22-µm filter. Adjust pH if necessary and pour filtered buffer into buffer reservoir.

    If the column is going to be used for a prolonged time, it is good practice to continuously deaerate the GF buffer by heating the buffer slightly with heat from a light bulb, or to use a degasser.

15. Assemble GF system according to manufacturer's instructions or Figure 8.3.1, placing detector and chart recorder (if used) in line but leaving column off line. Attach buffer reservoir to pump and purge pump with GF buffer with column off line.

    If the system is run by gravity flow alone, the risk of the column running dry may be eliminated by placing the end of the outlet tubing higher than the inlet adaptor or placing a loop of the tubing from the reservoir with its bend at a lower level than the outlet.

    Detection of salt may be accomplished in line by conductivity or off line by flame photometry or conductimetry. Colorimetric detection of proteins is readily done at 280 or 254 nm; the latter wavelength provides higher sensitivity. Refer to Chapter 3 for specific methods of protein detection.

16. Connect outlet of pump to column via the injection valve, run system with GF buffer at flow-rate settings to be used for separation, collect fractions with fraction collector, and note actual flow rate of the pump (e.g., by weighing collected fractions or measuring them with a graduated cylinder).

    If the concentration of the solute is detected by in-line monitoring, use of a two-channel recorder set at different sensitivities (e.g., 100% and 10% full-scale) is recommended to ensure that the peaks of unknown solutes are not off-scale.

    The actual flow rate of the system must be determined with the column installed because the flow rate from a peristaltic pump or similar device depends upon the back-pressure of the system.

Determine separating volume of column

17. Determine the void volume of the system (V_o) by running a void marker and obtaining the elution volume according to steps 20 to 23.

    The sample volume for determination of void volume and total liquid volume (steps 17 and 18) is not critical and may typically be 1% to 2% of the bed volume. However, correction of the calculated elution volume is needed if large sample volumes are applied (see annotation to step 23a). See Figure 8.3.2 for graphical depiction of void and total liquid volume.

    The markers for void volume and total liquid volume may be mixed and run together.

    Void volume (V_o) is typically 30% to 33% of the bed volume of a nonrigid (e.g., agarose) gel and 36% to 40% of the bed volume of a rigid (e.g., silica) gel.

BCU IASI CENTRAL UNIVERSITY LIBRARY

18. Determine the total liquid volume of the system (V_t) by running a total liquid volume marker and obtaining the elution volume according to steps 20 to 23.

Total liquid volume (V_t) is typically 85% to 95% of the geometric bed volume ($\pi r^2 L$, where r is the inner radius of the column) for a nonrigid gel and 70% to 80% of the bed volume for a rigid gel. Because solute-matrix interactions may occur, it is advisable to check that the total liquid volume is smaller than the bed volume and perhaps run several different total liquid volume markers to confirm the result.

19. Calculate the separating volume of the column (V_i) as $V_t - V_o$.

The separating volume (V_i) is the amount of liquid that will pass through the column from the point where large proteins begin to emerge to the point where small solutes are eluted. It is equal to the pore volume of the bed, which is typically 50% to 65% of the bed volume of a nonrigid gel and 30% to 50% of the bed volume of a rigid gel.

Prepare, apply, and chromatograph the sample

20. Dissolve sample to be desalted in GF buffer. Filter through a 0.22-µm protein-compatible filter.

If possible, let sample solution reach room temperature before being applied to column to reduce viscosity effects. This is more important for large sample volumes.

The influence of the sample volume on the resolution (see Critical Parameters) is related to the need for column efficiency (see Basic Protocol 2 and UNIT 8.1). Too large a sample volume may ruin the separation of closely eluting proteins in a separation where high efficiency and narrow peaks are required. This is of primary interest in protein fractionation (see Basic Protocol 2), whereas in desalting, conditions are chosen to yield maximum separation of the protein from contaminants. To this end, column efficiency is less important, and very large sample volumes may be applied. See Critical Parameters for guidelines on optimal sample concentration.

When GF is performed for analytical purposes, the sample concentration should be as small as possible within the sensitivity of the detection method employed. For preparative purposes, on the other hand, the concentration should be as high as possible. The limiting factor in the latter case is the viscosity of the sample divided by the viscosity of the GF buffer (relative viscosity). At relative viscosities >1.5, the hydrodynamic instability of the trailing edge of the zone may cause distortion of the

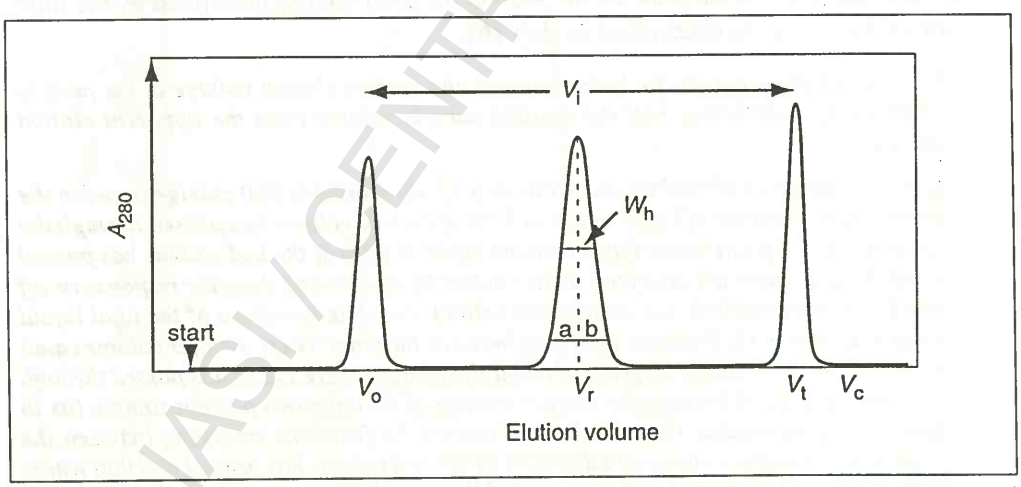


Figure 8.3.2 Chromatogram illustrating the separating volume (V_i) of gel filtration as determined by the void volume (V_o) and the total liquid volume (V_t). The total bed volume (V_c) is equal to the total liquid volume plus the matrix volume. Column efficiency is calculated as $N = 5.54(V_r/W_h)^2$, where N is the efficiency of the column in terms of number of theoretical plates, V_r is the elution volume of the peak, and W_h is the peak width at half peak height. The symmetry of the peak is calculated at 10% peak height as b/a, where b is the width of the tailing part of the peak and a is the width of the leading part of the peak.

UNIVERSITY LIBRARY

protein zone (resulting from invasion of less viscous eluant into the more viscous zone; called "viscous fingering"), which will ruin the separation. It is generally safe to use protein concentrations up to 50 mg/ml and in some cases up to 70 mg/ml, whereas concentrations of 100 mg/ml often result in viscous fingering. However, when large sample volumes are applied—e.g., in desalting—the low dilution factor will result in a high concentration of protein in the peak, hence lower concentrations than these may be required to avoid viscosity effects.

21. Open outlet from column, turn on pump, and let two bed volumes of GF buffer pass through the column. Turn on detector and recorder and allow baseline to stabilize.

22. Load sample applicator with the appropriate volume of sample for desalting, not exceeding the separating volume determined in step 19. Switch the sample application valve to the load position and put a start mark on the chart recorder paper (or start a stopwatch if no chart recorder is used).

    Smaller sample volumes are used for protein fractionation than for desalting, and much smaller sample volumes are used for analysis.

    Before injecting, flush injector loop with excess sample to avoid trapping air bubbles, which otherwise may end up in the top of the column.

    If the separating volume of the column has not been determined, the maximum sample volume may be estimated as ~50% of the bed volume for nonrigid media and ~30% of the bed volume for rigid media. For a first attempt at desalting, 50% of these volumes should be applied (i.e., 25% and 15% of the bed volume for nonrigid and rigid media, respectively).

    If an open-bed column is used, gently add GF buffer to the top of the column with a Pasteur pipet or similar device, then open the outlet from the column and continue adding GF buffer until two bed volumes of buffer have passed through. Let the last portion of buffer penetrate the gel without the gel surface becoming dry; this is facilitated by using a top filter on the bed. Gently apply the sample to the clear gel surface or the top filter using a micropipet (or a Pasteur pipet if large volumes of sample are to be desalted).

23a. *To determine elution volume:* Pass GF buffer through the system at the appropriate flow rate and follow detector response with recorder while collecting fractions. Construct a chromatogram and calculate elution volume as the time from the start point to the apex of the peak for the protein (or other solute) multiplied by the flow rate of the pump (as determined in step 16).

    If large sample volumes are being applied, the correct elution volume of the peak is obtained by subtracting half the applied sample volume from the apparent elution volume.

    For determination of void volume as in step 17, the fractions that emerge between the point where a volume of buffer equal to 25% of the bed volume has passed through the column and the point where buffer volume equal to 50% of the bed volume has passed through the column are analyzed (either in line by monitoring detector response or off line by another method; see annotation below). For determination of the total liquid volume as in step 18, fractions emerging between the point where a buffer volume equal to 75% of the bed volume has passed through and that where 125% has passed through are analyzed. To determine the elution volume of an unknown protein sample (as in determining molecular size; see Basic Protocol 3), fractions emerging between the point where a buffer volume equal to 25% of the bed volume has passed and that where 125% has passed through are analyzed. Allowing 125% to pass through will permit the entire peak to be traced.

    If fractions are to be collected for subsequent off-line analysis to determine where elution occurs (e.g., by photometry at 280 nm to detect Blue Dextran 2000 or proteins, or conductivity to trace the salt peak), start and stop the fraction collector to sample fractions at the volumes stated above. The fraction size should be large enough to

PCU IASLIC CENTRAL UNIVERSITY LIBRARY

permit accurate assay of the solute of interest; often 1 ml or 1/50 of the bed volume (whichever is smallest) is a suitable fraction size. The prefraction (i.e., the amount of liquid eluted from the point of sample application to the start of fraction collection for analysis) should be sampled in a vessel of known weight. The elution volume is calculated from the volume of the fractions by adding half the volume of the peak fraction and the volume of the prefraction, then subtracting half the sample volume. For dilute aqueous buffers (e.g., 0.5 M) a density of 1 may be assumed for converting weight to volume. If the experiment will take a long period of time, it is advisable to minimize evaporation by sealing the fraction cups with Parafilm.

If an open-bed (gravity flow) column is used, carefully add a volume of GF buffer that corresponds to the void volume of the column reduced by half the sample volume—unless the sample volume is <4% of the bed volume, in which case add 1 void volume of buffer. If the column is operating correctly, the effluent collected should contain no protein. Add GF buffer in ≤1 ml increments if the column has a bed volume <50 ml, collect fractions, assay for proteins, and calculate elution volumes as described above.

23b. *To desalt a protein:* Pass a volume of GF buffer equal to the void volume reduced by half the applied sample volume through the column, and collect a waste fraction. Place a vessel that can hold 1.5 times the applied sample volume under the column, then add a volume of GF buffer equal to the applied sample volume. Collect the desalted protein.

Depending upon the efficiency of the desalting system, these quantities may have to be adjusted (i.e., the volume collected in the second fraction may need to be decreased in order to preclude contamination from substances that elute later). If system effects (e.g., mixing after the column) are substantial or the column is badly packed (e.g., contains cracks in the bed), it will be necessary to pass more GF buffer through the column than the void volume minus half the sample volume before sampling the protein (see Critical Parameters). Figure 8.3.3 shows a sample chromatogram of optimized desalting of 400 ml hemoglobin.

The desalting step may be continuously monitored by simultaneous in-line detection

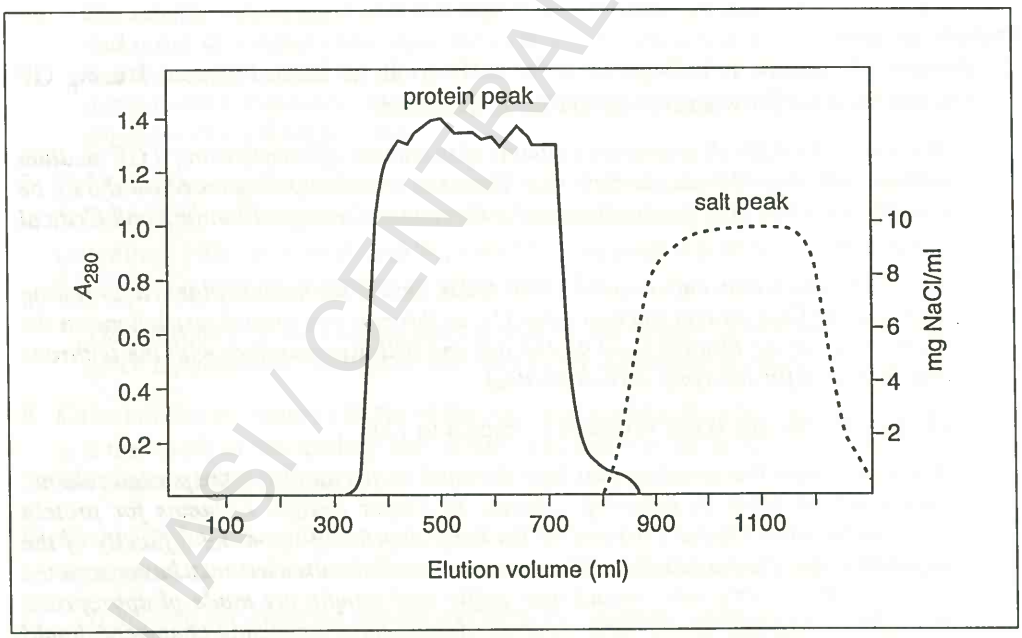


Figure 8.3.3 Chromatogram illustrating desalting of proteins by gel filtration. A 4 × 85–cm column packed with Sephadex G-25 is used. The sample consists of 400 ml hemoglobin (protein peak; solid line) in NaCl (salt peak; broken line). Note that the sample volume is close to the pore volume of the packed matrix, i.e., 490 ml. No correction for sample volume has been made. Calculation from the figure yields V_0 = 560 ml – 200 ml = 360 ml, which is 31% of the geometric column volume (V_c). Reproduced from Flodin (1961) with permission of the *Journal of Chromatography.*

UNIVERSITY LIBRARY / CENTRAL UNIVERSITY OF IASI

of the protein zone by absorbance at 280 nm and in-line detection of the salt peak by conductivity.

24. Wash column with ≥1 column volume of GF buffer containing an antibacterial agent. Close column outlet and store column.

    See recipe for GF buffer for acceptable antibacterial agents.

BASIC PROTOCOL 2

PROTEIN FRACTIONATION

Protein fractionation refers to the separation of proteins of similar molecular size for purification purposes. The objective may be to separate dimer and higher aggregates from a pharmaceutically active protein or peptide. Requirements for the column and system are therefore substantially higher for protein fractionation than for desalting.

The same materials and methodology described for desalting are used for protein fractionation. However, selection of matrix and column (see Strategic Planning) is more critical in this procedure, as are the quality of the column packing and control over experimental conditions (e.g., temperature and flow rate). The protocol includes steps for determining the column efficiency (see also *UNIT 8.1*) to evaluate the fitness of a packed column for use in protein fractionation.

Materials

GF matrix (see Table 8.3.4) or prepacked GF column (see Table 8.3.6) with appropriate selectivity for protein of interest (also see Strategic Planning and Critical Parameters)
GF fractionation buffer (see recipe)
Colored marker: 0.2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B_{12}
Low-molecular-weight marker (e.g., 5 mg/ml acetone or 2 M sodium chloride)
Protein sample to be fractionated
GF chromatography column (see Critical Parameters and Table 8.3.5)

Prepare column

1. Prepare GF matrix as in steps 1a to 3a or 1b to 3b of Basic Protocol 1 using GF fractionation buffer wherever GF buffer is indicated.

   It is generally difficult to prepare columns of maximum efficiency using a GF medium with a small (e.g., 10-μm) particle size. The need to scale up the procedure should be considered when selecting medium and column (see Strategic Planning and Critical Parameters).

   The GF fractionation buffer must be selected in light of the method planned for testing purity of the final protein fraction (step 11), as this may put special restrictions on the buffer (e.g., azide absorbs light at 254 nm and will also interfere with the anthrone reaction used for assaying carbohydrates).

2. Pack the column (see Basic Protocol 1, steps 4 to 12).

   Because protein fractionation puts high demands on the quality of the packed column, care must be taken in selecting columns of proper design. Columns for protein fractionation are typically 30 to 100 cm long, depending upon the difficulty of the separation and the particle size of the medium used. Columns used must be constructed so that parts coming into contact with buffer and sample are made of appropriate materials. Long tubings and other sources of dead space or mixing chambers should be avoided. The pressure resistance of the column should be compatible with the expected pressure drop of the bed. See Strategic Planning and Critical Parameters for details concerning column selection.

3. Check quality of column packing visually (see Basic Protocol 1, step 13).

BCU IASH LIBRARY CENTRAL UNIVERSITY

4. Assemble and test the system (see Basic Protocol 1, steps 14 to 16).

Place system and column in an environment not subject to temperature variations (e.g., the column should not be exposed to direct sunlight). If the system is to be used in a cold room, the influence of temperature on viscosity must be considered. Zone broadening of proteins and the pressure drop of the system are proportional to the viscosity of the buffer, and the flow rate will have to be adjusted accordingly.

Successful, reproducible protein fractionation requires a constant flow rate. This may be achieved using a peristaltic or high-precision pump whose flow rate is regularly checked. Continuous monitoring of the effluent is necessary to ensure proper sampling of the fraction of interest. Detection may be carried out at 280 or 254 nm to trace proteins containing aromatic amino acids, or at lower wavelength (e.g., 214 or 206 nm) to measure absorption by peptide bonds. Other parts of a protein fractionation system that might be of use include a fraction collector, (although in the simplest case it is possible to collect the fraction of interest by using a switch valve), and a chromatography controller, which permits automatic fractionation of the protein of interest.

Evaluate column

5. Chromatograph a colored marker (2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B_{12}; see Basic Protocol 1, steps 20 to 23) and ascertain that the zone produced is horizontal and sharp.

A volume of marker solution 1% to 4% of the bed volume should be applied. The zone produced by a colored sample must be horizontal and sharp for packing to be considered acceptable.

If the supplier recommends a different method than this for evaluating column packing, that method should ordinarily be used.

6. Chromatograph a low-molecular-weight marker (e.g, 5 mg/ml acetone or 2 M sodium chloride; see Basic Protocol 1, steps 20 to 23) and construct a chromatogram to determine the column efficiency.

The sample volume applied in this step may contribute considerably to the peak width and result in a falsely low value for column efficiency. It is therefore recommended that the sample volume not exceed 0.9%, 0.5%, and 0.4% of the bed volume for columns packed with GF matrices having average particle diameter of 100 μm, 30 μm, and 10 μm respectively (Hagel, 1985).

7. Calculate the number of theoretical plates per column according to the equation $N = 5.54(V_r/W_h)^2$, where N = number of theoretical plates per column, V_r = elution (retention) volume of peak, and W_h = width of the peak at half peak height. See Figure 8.3.2 for a graphical depiction of these variables.

The efficiency of the column corresponds to the number of theoretical plates under given experimental conditions. See UNIT 8.1 for discussion of column efficiency.

8. Calculate the asymmetry factor of the peak according to the equation $A_s = (b/a)$, where a is the width of the leading part of the peak and b is the width of the tailing part of the peak at 10% peak height.

9. Compare the plate number and asymmetry factor obtained for the column with the acceptance limits for these parameters in the manufacturer's documentation.

If no data for expected column performance is given, calculate the plate height (H) according to the equation $H = L/N$, where L is the length of the packed bed and N is the number of theoretical plates calculated in step 7. As a rule of thumb, H should be about three times the average particle diameter of the matrix for a good column and twice the average particle diameter for an excellent column. Plate heights of five times the average particle diameter or more are generally not accepted for high-resolution chromatography. The asymmetry factor should be 1.0 ± 0.2.

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

Table 8.3.7 Practical Velocities for Protein Fractionation or Size Determination[a]

Solute molecular weight (Daltons)	Nominal eluant velocities (cm/hr) at a support particle size of (μm):						
	2	5	10	15	30	50	100
1,000	507	203	101	68	34	20	10
5,000	293	117	59	39	20	12	6
10,000	234	94	47	31	16	9	5
15,000	195	78	39	26	13	8	4
50,000	137	55	27	18	9	6	3
75,000	121	48	24	16	8	5	2
100,000	109	44	22	15	7	4	2
200,000	86	34	17	11	6	3	2
300,000	76	30	15	10	5	3	2

[a]Adapted from Hagel (1989) with permission from VCH Publishers.

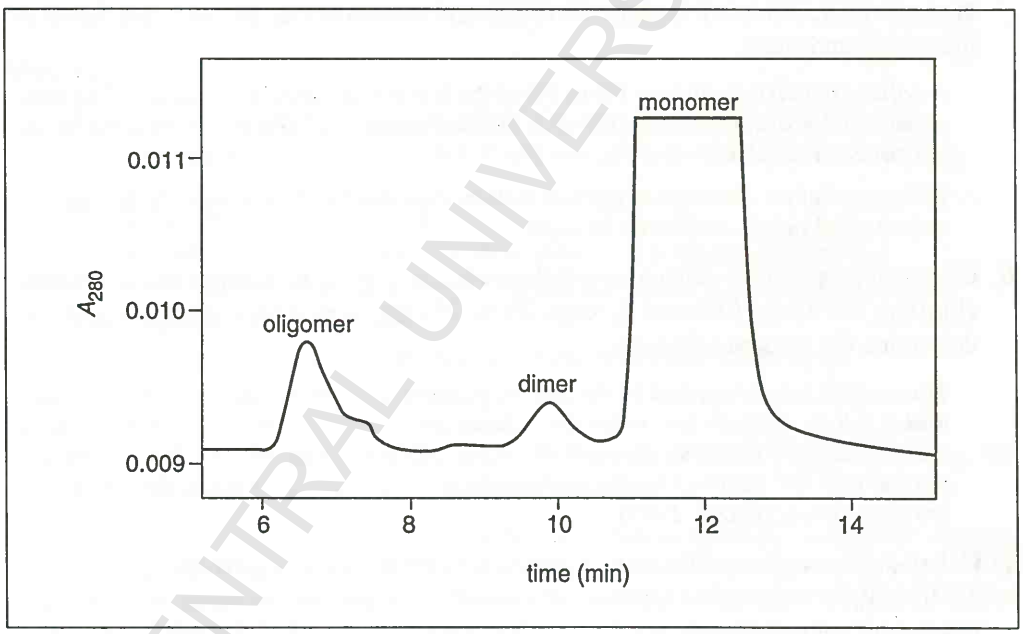


Figure 8.3.4 Chromatogram illustrating protein fractionation by gel filtration. A Superdex 75 HR 10/30 column is used. The sample consists of recombinant human growth hormone. This figure is an enlarged section of the original chromatogram displayed to illustrate complete resolution between monomer and dimer. Reproduced from Hagel (1993) with permission of the publisher. Courtesy of B. Pavlu, Kabi-Pharmacia Peptide Hormones, and H. Lundström, Pharmacia Biotech.

The flow rate used will influence the broadening of the solute peak; suggested flow rates for GF matrices of different particle size are given in Table 8.3.7.

Perform chromatography and evaluate results

10. Dissolve, apply and chromatograph protein sample to be fractionated (see Basic Protocol 1, steps 20 to 23).

Start by applying a small fraction of the sample to get a rough idea of the separation and to calculate optimum conditions for the actual run. A sample volume of ~1% to 4% of the bed volume has been found to be optimal for protein fractionation using a medium of 30-μm particle size, depending upon the processing rate (i.e., amount of substance processed per unit time) of the system (Hagel et al., 1989). As a general rule,

BCU IASI / CENTRAL UNIVERSITY LIBRARY

a sample volume of 2% of the bed volume can be applied. A chromatogram illustrating protein fractionation where dimers and higher aggregates are separated from the target monomeric form of the protein is given in Figure 8.3.4.

Optimum sample volume depends upon many experimental factors; discussion of optimization strategies have been published (Hagel et al., 1989; Hagel and Janson, 1992). In general, the larger the bead size used, the larger the sample volume that may be applied before the initial resolution is lost. This is due to the fact that smaller bead sizes normally result in more efficient columns that produce narrower peaks; hence the influence of sample volume on peak width is more detrimental.

11. Check the purity of the collected fraction(s).

Suitable methods of determining purity include HPLC and electrophoresis (see Chapter 10 for electrophoresis).

DETERMINATION OF MOLECULAR SIZE

BASIC PROTOCOL 3

The third important application of gel filtration is in determining molecular size. To accomplish this, a GF column is calibrated with reference samples of known molecular size (see Support Protocol). Gel filtration is then carried out on an unknown protein solute as in Basic Protocols 1 and 2, and the molecular size of that solute is then inferred from the calibration curve, which relates elution volumes or distribution coefficients on the calibrated column to the molecular sizes of the different reference samples. If the relationship between the molecular size and mass is known, the mass of the solute may be calculated. However, because the shapes of biological molecules vary considerably (e.g., proteins may be spherical or ellipsoidal, or in some cases more or less rod-shaped), the calculation of molecular mass (M_r) from elution volume must be done with great care. If the shape of the molecule under study is not known, the experiment will yield only its estimated solvated size. The influence of solute shape on size may be exemplified by the fact that the radius of gyration (one estimate of size) is proportional to $M_r^{1/3}$ for spherical molecules (e.g., some proteins), to $M_r^{1/2}$ for molecules shaped like flexible coils (e.g., dextrans and denatured proteins), and to M_r for rod-shaped molecules (e.g., fibrous proteins and DNA fragments of intermediate length). Thus, gel filtration is very suitable for estimating the molecular size but generally not the molecular mass of unknown solutes.

Materials

GF matrix or prepacked GF column (see Table 8.3.4, Table 8.3.6, Strategic Planning, and Critical Parameters)
GF fractionation buffer (see recipe; note variations for size determination)
GF chromatography column (Table 8.3.5)
High-precision pump
Protein sample for size determination

1. Prepare gel, pack column, then assemble and test GF system (see Basic Protocol 1, steps 1 to 16) using GF fractionation buffer wherever GF buffer is indicated.

For high-precision analytical gel filtration, the temperature of the chromatographic bed is often kept constant by using a jacketed column and circulating water from a temperature-controlled water bath. This makes the instrument setup more complicated, and the need for this precaution must be assessed on a case-by-case basis (e.g., if the column is calibrated at every other run, as is possible with high-speed gel filtration, there is little need for temperature control).

A high-precision pump must be used to ensure accurate and constant flow rate, which is of utmost importance for achieving reliable results in molecular size determinations. In addition to a UV detector for determining protein concentration, a detector for viscosity and/or light scattering—e.g., one employing multiple-angle laser light scat-

BOURNEMOUTH UNIVERSITY LIBRARY

tering (MALLS)—may be used for assay of molecular size or (indirectly) molecular mass. A mass spectrometry detector—e.g., one employing electrospray ionization mass spectrometry (ESI-MS)—may be used for assay of molecular mass.

If denatured standards are to be used in the calibration to eliminate the influence of molecular shape (see Support Protocol), 6 M guanidine hydrochloride (see recipe for GF buffers in Reagents and Solutions) is used as the GF fractionation buffer in this and subsequent steps.

Once a column has been calibrated (see Support Protocol), it is important to check column performance by determining the elution volume of one of the reference samples used for calibration as part of the routine testing of the system.

2. Determine the void volume (V_o) and the total liquid volume (V_t) for the system (see Basic Protocol 1, steps 17 and 18).

3. Calibrate column (see Support Protocol). Check flow rate during calibration by sampling effluent and weighing fractions.

   Calibration of the column (and chromatography of the sample) may be carried out under denaturing conditions to eliminate the effects of molecular shape. Variations in procedure are described in the Support Protocol.

   There are no general restrictions on the flow rate as long as the protein of interest separates from contaminating solutes (ideally, baseline separation should be achieved). A high flow rate will not influence the qualitative information sought—i.e., the position of the peak maximum. Thus, qualitative information may be obtained by very fast gel filtration; however, the danger of shear effects at high flow rates on elongated solutes must be considered. If a high-precision pump is used, there is less need to check the flow rate for every run (though in the author's experience this is a good practice because any pump may malfunction). If constant flow rate can be taken for granted, time may be substituted for volume in the calculations.

   For both the calibration (see Support Protocol) and the run to determine molecular size of the unknown sample, the elution volume should generally be corrected for the contribution from the sample volume (see Basic Protocol 1, step 23a). However, if the sample volume is kept constant throughout the entire investigation it may arbitrarily be regarded as a part of the dead volume of the system, and no correction is necessary.

4. Apply, elute, and chromatograph unknown sample (see Basic Protocol 1, steps 20 to 23). Prepare a chromatogram and calculate elution volume. Check flow rate by sampling the effluent and weighing fractions.

   The same applied sample volume should be used for the unknown sample as for the size standards (see Support Protocol).

   If denatured standards were used in the calibration to eliminate the effects of molecular shape (see Support Protocol) the sample must be denatured. Elution is then carried out using 6 M guanidine hydrochloride as the GF buffer.

   Often the calculations are performed on a time basis using an integrator or chromatography software (Pharmacia Biotech).

5. Calculate molecular size of unknown sample components using the calibration graph constructed in the support protocol.

   If two different detectors were used for the reference standards and the sample (e.g., a refractive index detector for calibration with dextran standards and a UV detector for the protein sample), a correction factor for the difference in volume between the two detectors must be used in the calculations.

   The estimate of molecular size will be correct only if there are no sample-matrix or sample-sample interactions (in fact, gel filtration has been used to study these self-association processes). The size estimate will be an apparent gel-filtration size

BCU IASI / BIBLIOTECA UNIVERSITARA

which, together with data about the molecular mass (e.g., from mass spectrometry), will yield information about molecular shape.

6. Continue to run GF buffer through column until the next run is to be made.

It is important to continue flushing the column to maintain column characteristics and avoid the need for recalibration (also see Critical Parameters). If the next run is not to be made for a long time, the flow rate may be reduced and buffer recirculated through a filter to the buffer reservoir (also see Critical Parameters). Before reusing the column after a long period of rest, the validity of the calibration curve should be checked by running two or three reference samples covering the range of interest and calculating their molecular size from the old calibration curve. The apparent size should agree with the assigned size within ±5% error.

COLUMN CALIBRATION

SUPPORT PROTOCOL

Molecular size standards for column calibration should ideally be of the same shape and type as the molecules under study (e.g., globular proteins, fibrous proteins, or peptides). In cases where this requirement cannot be met, the column can be calibrated with other suitable reference substances of known size (e.g., dextrans). An apparent size may thus be assigned to the unknown and the actual size then inferred from the relationship between the shape of the different molecules. It is thus possible to use dextran as a reference for the estimation of Stokes radius for globular proteins (Hagel, 1993). However, a general shape parameter governing separation of all types of molecules by gel filtration has not yet been found (Dubin et al., 1990).

When protein standards are used, the calibration curve will reflect small natural variations in the molecular shapes of proteins, which will decrease the precision of the curve. However, the procedure described here is very simple and will readily yield an apparent molecular mass or molecular weight, which in many cases is quite satisfactory.

The influence of protein shape on the calibration curve may be eliminated by disruption of noncovalent bonds with denaturing solvents (e.g., guanidine hydrochloride). Disulfide bonds are broken by reductive cleavage with a reagent such as DTT, and reoxidation is prevented by carboxymethylation. This treatment converts the shape of sample and reference proteins to random coils. The physical size of a random coil is larger than that of a compact molecule of equal mass; therefore a more porous gel is needed to separate the former species. Detergents may also be used to denature proteins. However, it must be noted that a protein-detergent complex is larger than the protein and that ionic detergents will introduce charged groups. Also, information about size or molecular mass obtained using detergent-denatured proteins may not be relevant to the native protein. As these types of treatment are frequently used to prepare proteins for analytical electrophoresis, the reader is referred to Chapter 10 for further details.

The influence of variation in shape on the calibration curve may also be eliminated by using as molecular size standards reference substances from a homologous class of nonprotein polymers—e.g., dextran or polyethylene glycol. Dextran has been used extensively as reference substance for calibration of gel-filtration columns. Calibration can be done by running a number of narrow fractions or one sample of broad size distribution.

Additional Materials (also see Basic Protocol 3)

Calibration standards (see recipe and Table 8.3.8)
6 M guanidine hydrochloride (see recipe for GF buffer)
In-line refractive-index detector

BU JAST CENTRAL UNIVERSITY LIBRARY

For calibration with native proteins

1a. Calibrate the column using protein standards (see Table 8.3.8) as in steps 20 to 23 of Basic Protocol 1 and construct a chromatogram.

It is practical to mix a number of the calibration standards and chromatograph them simultaneously so that the calibration may be completed in a few runs (see example in Figure 8.3.5).

2a. Plot the elution volume (V_r; i.e., the volume on the chromatogram from the injection point to the apex of the peak for a particular size standard) versus the logarithm of the molecular size, $\log_{10}R$, where R is the molecular radius of the calibration standard.

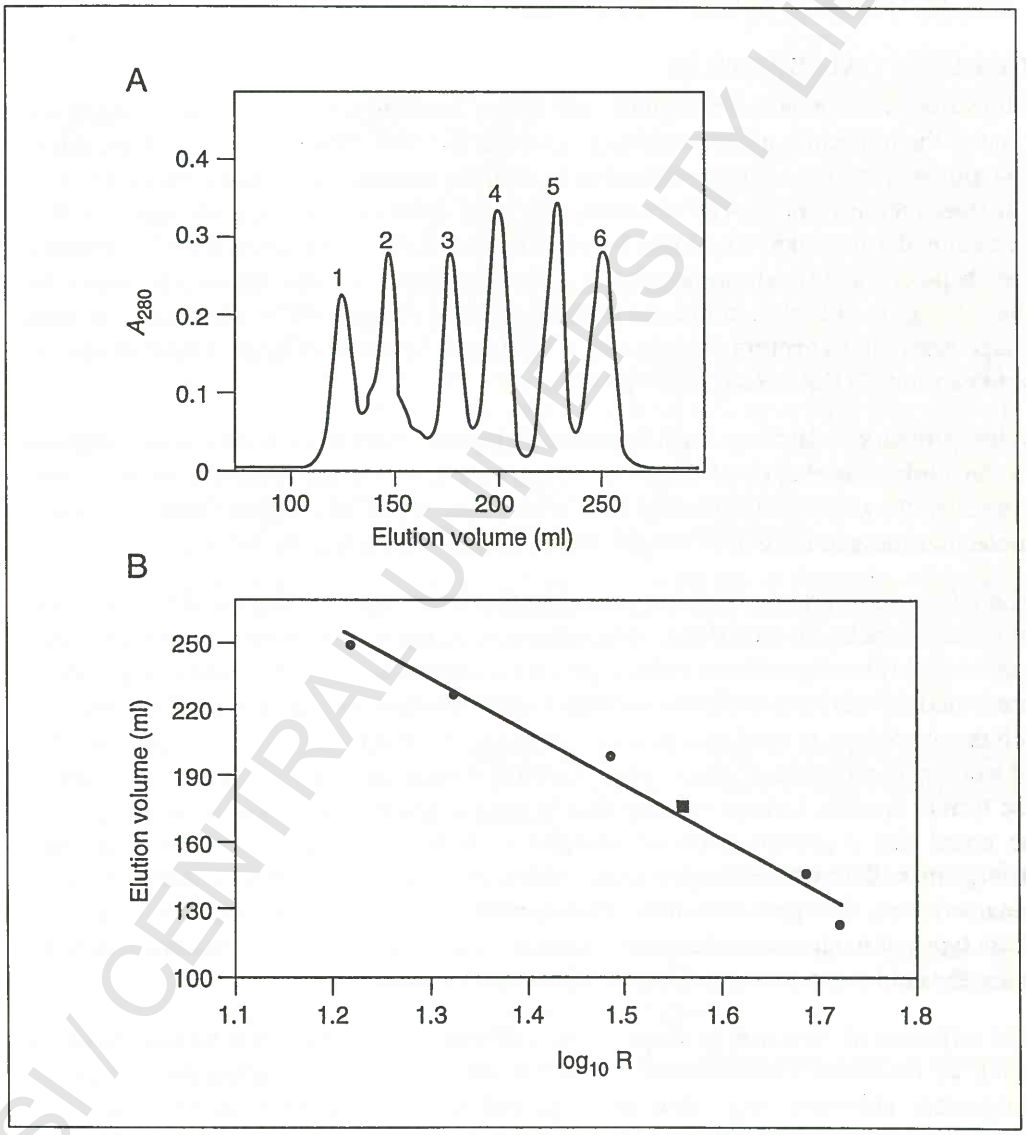


Figure 8.3.5 Calibration of GF column for determination of protein size. A K 26/70 column packed with Sephadex G-200 SF is used. (**A**) Chromatogram obtained for six size standards: 1, catalase; 2, aldolase; 3, bovine serum albumin; 4, ovalbumin; 5, chymotrypsinogen A; 6, ribonuclease A. (**B**) Calibration curve obtained by plotting the logarithms of the molecular radii ($\log_{10}R$) of the six standards (*x* axis) versus the elution volumes for peaks 1, 2, 4, 5, and 6 from panel A (filled circles). The curve was used for evaluating the size of bovine serum albumin (filled square; elution volume corresponds to peak 3 in panel A), giving an apparent size of 34.8 Å. The excellent agreement with the nominal molecular size value of 35.5 Å may be incidental. Repeating the procedure for peak 4 in panel A showed that a deviation of ~3 Å from the nominal value may be expected for proteins of globular shape. Panel A reproduced from Pharmacia Biotech (1991) with permission from the publisher.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 8.3.8 gives the molecular radii of commonly used calibration standards. A sigmoid curve should be obtained that is approximated by a straight line in the middle section (see Fig. 8.3.5).

If the sample volumes are not held constant, the elution volume must be normalized by setting the injection point at half injected volume.

The calibration curve is often made system-independent by plotting distribution coefficient (K_D) versus $\log_{10}R$. The distribution coefficient is calculated as

$$K_D = (V_e - V_o)/(V_t - V_o),$$

where V_e is the elution volume for a particular peak, V_o is the void volume for the column, and V_t is the total liquid volume for the column.

For substances belonging to a homologous series (e.g., dextrans), it is meaningful to plot elution volume versus the logarithm of the molecular mass ($\log_{10}M_r$). Where substances of different chemical nature are being used to calibrate the column, the preferable way to calibrate the column is by using the hydrodynamic volume V_h. This is calculated according to the equation $V_h = (\eta)(M_r)/N\nu$, where η is the intrinsic viscosity, M_r is the relative molecular mass, N is Avogadros number (6.02×10^{23} molecules per mole), and ν is a shape factor.

It is wise to check the accuracy of the calibration by deleting one point at a time from the calibration data set and calculating the size or mass of the protein represented by that point from the calibration curve obtained with the remaining data. The average absolute difference from the nominal value indicates the accuracy of the calibration (see Fig. 8.3.5).

For calibration with denatured standards

1b. Denature calibration standards according to a standard denaturing protocol.

2b. Carry out steps 1a and 2a above, using a column equilibrated in 6 M guanidine hydrochloride and using 6 M guanidine hydrochloride as the GF buffer in the cross-referenced steps of Basic Protocol 1.

Where denatured standards are used, denaturing conditions must be maintained in chromatography of the sample (see Basic Protocol 3).

For calibration with nonprotein polymer standards

1c. Using an in-line refractive index detector, run system until a stable baseline is obtained. Chromatograph a number of dextran calibration standards (see Basic Protocol 1, steps 20 to 23) that will give calibration points surrounding the protein under study, to obtain their elution volumes or distribution coefficients.

The proper number and size of calibration standards to be run may be quickly found by running first the unknown sample and then the necessary calibration samples that will be eluted close to the unknown protein.

Apply each standard individually or as a mixture of two dextrans of high and low molecular mass.

Using a two-channel recorder set at two different sensitivities may prevent the peak from running off-scale.

2c. Plot the elution volume (V_e) or distribution coefficient (K_D) versus the logarithm of the viscosity radius of dextran ($\log_{10}R_{vis}$).

The viscosity radius of dextran, R_{vis}, is related to the molecular mass (M) according to the equation $R_{vis} = 0.271 \times M^{0.498}$. Because dextran is a polydisperse polymer, the value of M used here is the molecular mass of dextran that corresponds to the peak apex, sometimes denoted M_p.

When two different detectors are used for calibration and chromatography of the sample—e.g., a refractive index detector for dextran standards and a UV detector for

BOKU-VIBT / CENTRAL UNIVERSITY LIBRARY

the protein sample—the difference in volume between the two detectors must be taken into account in calculation of elution volume.

Another procedure where the entire separation range of interest was calibrated in one run, using an integral calibration method, has been described (Hagel and Andersson, 1991; Hagel, 1993). Unfortunately, this simple and powerful method cannot yet be performed on routine basis because it requires data for the molecular mass distribution of the sample, and this information is, at the moment, not readily available in most cases.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

GF buffers

For desalting: Use a volatile buffer—e.g., 0.05 M ammonium hydroxide, 0.05 M acetic acid, 0.05 M ammonium acetate or 0.5 M ammonium bicarbonate—unless the step is for buffer exchange, in which case use the desired final buffer.

Distilled (Milli-Q purified or equivalent) water may be preferred to buffer as an eluent in some cases. This will yield adequate results with noncharged solutes; however, ionic interactions with the small amount of charged groups on the gel surface may disturb the purification of slightly charged molecules when water alone is used. Addition of a volatile bacteriostatic agent—e.g., chlorhexidine—may sometimes provide enough ionic strength to suppress such interactions.

For protein fractionation: Use 0.05 M sodium or potassium phosphate (pK_a = 2.1, 7.2, 12.4), 0.05 M Tris·Cl (pK_a = 8.1), or 0.05 M sodium or potassium acetate (pK_a = 4.8).

The buffering capacity of these buffers is adequate at a pH within ±0.5 units of the pK_a value. If the protein is to be freeze-dried, use of a volatile buffer or water is advantageous.

It is sometimes necessary to add 0.1 M sodium chloride to completely eliminate ionic interactions. If parts of the system that come into contact with the buffer are composed of steel, the sodium chloride should be replaced with 0.05 M sodium sulfate to prevent halide-induced corrosion. In some cases the ionic strength of the buffer will affect the sample, e.g., by disrupting aggregates. This effect can be eliminated only by changing the ionic strength.

For determining molecular size: Use the same buffers as for protein fractionation but select one that will preclude solute-matrix interactions (see Support Protocol). If the column is to be used for a long period, add preservatives to prevent bacterial growth. Make sure the buffer is compatible with the detection step (e.g., use a buffer with low salt content if mass spectrometry will be used for detection).

If calibration is to be conducted under denaturing conditions, 6 M guanidine hydrochloride is used as the GF buffer. Guanidine hydrochloride is a better denaturing agent than urea.

For all buffers: Add antimicrobial agent if buffer will be used for a long time, adjust the pH of the buffer, filter through a 0.22-µm filter, then recheck pH and readjust if necessary.

It is also recommended to degas the buffer prior to prolonged use. It is common practice to prepare fresh buffer once a week, but if an antimicrobial agent has been added it may be stored for 1 month or more. Antimicrobial agents used for GF buffers include ethanol added to a concentration of 20% (v/v), sodium azide added to a concentration of 0.02% to 0.05% (w/v), or chlorhexidine added to a concentration of 0.002% (w/v). Ethanol is not recommended for use with Sephadex gels as organic solvents cause the bed to shrink when used with this matrix.

BCU IASI CENTRAL UNIVERSITY LIBRARY

IMPORTANT NOTE: *Azide is a very potent antimicrobial agent and therefore widely used. However, handling of azide, especially the dry powder, requires special attention; please refer to information from the supplier for safety precautions. Frequent exposure to the dry powder may be avoided by preparing a stock solution (e.g., 20% w/v) from which aliquots (e.g., 1 ml azide stock solution to 1 liter GF buffer) may be drawn conveniently with a micropipet. Azide may form explosive salts in lead-pipe waste disposal systems; collect the effluent in waste bottles.*

Calibration standards

Table 8.3.8 provides a compilation of commercially available molecular size calibration standards and their suppliers. Some of these are available as kits (e.g., from Pharmacia Biotech). Protein standards should be prepared in the GF buffer that will be used in the molecular size determination. To prolong the life of the column, samples should be filtered through a protein-compatible filter prior to injection. The exact concentration to be prepared depends on the additives present (e.g., sucrose as stabilizer), detector sensitivity, and the zone broadening of the column. Concentrations of proteins in the range from 1 to 5 mg/ml are frequently used for calibration. It is possible to mix several standards to calibrate the separation range of interest in a few runs (it is seldom necessary to calibrate the entire separation range of the column).

Table 8.3.8 also lists some dextran fractions that are useful for calibrating the column in hydrodynamic volume, which makes comparing solutes of different shape easier.

Total liquid volume marker

5 mg/ml acetone, 10 mg/ml sodium azide, 0.1 mg/ml cytidine, 10 mg/ml sodium nitrate, and 2 M sodium chloride have been used for determination of total volume. A negative pulse from injection of water can also be used. In cases where interactions between matrix and solute must be eliminated, deuterium oxide can be used.

Acetone and cytidine are detected by UV absorbance at 280 nm, azide at 254 nm, and nitrate at 254 nm. Sodium chloride can be detected by flame photometry, conductivity, a sodium or chloride electrode, or precipitation of chloride with silver nitrate solution. A negative pulse from injection of water is detected by measuring the absorbance or refractive index of the buffer or by monitoring the conductivity of the effluent. Deuterium oxide is detected with a refractive-index detector.

Void marker

A suitable substance for determining void volume is a protein of the correct size (not too large and not too small), which may be selected by running a series of large proteins and observing which of them start to permeate the matrix. Other substances that have been used include 2 mg/ml Blue Dextran 2000 and DNA.

Blue Dextran 2000 has the advantage of being visible to the eye. Disadvantages are that small amounts of free dye may stick to the gel and a broad permeating peak may be visible in addition to the sharp excluded peak if the pore size of the gel is large enough (e.g., as with Sepharose 4B). This occurs because of the polydispersity of dextran. If the molecule is too large, so-called secondary exclusion effects, in which the sample is excluded from the narrow space where beads intersect, will yield a falsely low void volume. However, this phenomenon is only important in high-precision gel filtration and rarely has to be considered in desalting.

BCU IAS CENTRAL UNIVERSITY LIBRARY

Table 8.3.8 Molecular Size Standards for Calibration of Gel Filtration Columns

Standard	Molecular mass (M_r, Daltons)	$\log_{10}(M_r)$	Molecular size (radius, Å)[a]	Supplier[a]
Proteins[b]				
Thyroglobulin, bovine thyroid	669,000	5.825	85.0	PB
Ferritin, horse spleen	440,000	5.643	61.0	PB
Catalase, bovine liver	232,000	5.365	52.2	PB
Gamma globulin, bovine	158,000	5.199	—	BR
Aldolase, rabbit muscle	158,000	5.199	48.1	PB
Transferrin	81,000	4.908		SI
Albumin, bovine serum	67,000	4.826	35.5	PB
Ovalbumin, chicken	44,000	4.643	—	BR
Ovalbumin, hen egg	43,000	4.633	30.5	PB
Chymotrypsinogen A, bovine pancreas	25,000	4.398	20.9	PB
Myoglobin, equine	17,500	4.243	—	BR
Ribonuclease A, bovine pancreas	13,700	4.137	16.4	PB
Non-protein standard				
Vitamin B_{12}	1,350	3.130	—	BR
Dextrans[c]				
Dextran T-10	8,100		23	PB
Dextran T-40	23,600		40	PB
Dextran T-70	33,000		50	PB
Dextran T-500	370,000		150	PB
Dextran 1	1,080		9	PC
Dextran 5	4,440		18	PC
Dextran 12	9,980		26	PC
Dextran 25	21,400		39	PC
Dextran 50	43,500		55.3	PC
Dextran 80	66,700		68.4	PC
Dextran 150	123,600		93.1	PC

[a]Abbreviations and symbols: — Not stated by manufacturer; BR, Bio-Rad; PB, Pharmacia Biotech; PC, Pharmacosmos; SI, Sigma. Addresses and phone numbers of suppliers are provided in the *SUPPLIERS APPENDIX*.

[b]The size of proteins may vary according to the source of information and the methodology used for determination—e.g., a value of 67.1 Å has also been assigned to ferritin (de Haen, 1987).

[c]Dextran is a polydisperse polymer and values given here—e.g., peak molecular mass—may vary from lot to lot. In cases where peak molecular mass is not stated by the supplier, the column is calibrated as elution volume versus $(M_w M_n)^{1/2}$, where M_w is the weight-average molecular mass and M_n is the average molecular mass. The molecular size of dextran is calculated according to the equation $R_{vis} = 0.2 > 1 \times M_{0.498}$ where R_{vis} is the viscosity radius and M is the molecular mass (Hagel, 1989).

COMMENTARY

Background Information

Although separation of molecules according to size had been noticed in the mid-1950s (Lindqvist and Storgårds, 1955; Lathe and Ruthven, 1956), the discovery of gel filtration (GF) may be attributed to Porath and Flodin (1959), who realized the potential of the technique and were the first to report using it to desalt proteins. Their discovery provided the basis for a new separation technology that rapidly gained widespread acceptance owing to the commercial availability of a suitable cross-linked gel-filtration matrix—Sephadex—introduced in 1959. Remarkably enough, after 35 years this matrix is still one of the premier choices for desalting of proteins (see Table 8.3.1). The use of gel filtration for determination of molecular size of proteins was described later (Andrews, 1970). Gel filtration was also rapidly established as an important tool for

RCU IASI / CENTRAL UNIVERSITY LIBRARY

determining the molecular mass distribution of polymers in aqueous (Granath and Flodin, 1961) and organic solutions (Moore, 1964), for which suitable methods were lacking at that time. Moore used cross-linked polystyrene–based media for his application, naming the technique gel-permeation chromatography. Thus, within a few years after the discovery of gel filtration, applications of the technique basically covered the present range.

During the following years, the focus was directed on investigating the mechanism behind gel filtration. Theories suggesting that the separation was regulated exclusively by diffusion of molecules were soon replaced by the explanation, accepted today, that the process is driven by loss of conformational entropy when free permeation of the matrix by a solute is hindered by steric factors (Casassa, 1967). Of course, diffusion affects peak width and thus also the degree of separation, but diffusion does not influence the elution volume of a solute. The definite impact of steric exclusion on the separation has led to the proposal to rename the technique "steric exclusion chromatography" (ASTM, 1978), but this term has not gained wide acceptance. Another generic designation sometimes used in the literature is size-exclusion chromatography (SEC). It is important to realize that all of these designations refer to the same basic separation principle, which the majority of protein chemists refer to as gel filtration.

Development of gel-filtration media was focused for a long period on natural polymers or hydrophilic polymers. The late 1970s saw the introduction of silica-based materials, which combined rigidity and small particle size to enable fast gel filtration. A few years later, natural polymers of enhanced rigidity were introduced, which made it possible to pack columns with matrices of 10-μm bead size. To distinguish these media from the traditional type, the designation of high-performance size-exclusion chromatography (HPSEC) was introduced. However, it has been recently pointed out that the advantage of HPSEC media lies not in performance (i.e., traditional media packed in long columns may have equal or even better resolving capacity) but in speed of operation (Hagel, 1992). In addition, none of these materials have the outstanding selectivity shown by some of the traditional matrices (e.g., Sephadex and Bio-Gel P). Therefore, development of matrices for gel filtration has now been directed toward increasing their selectivity, which is done by decreasing the apparent width of the pore-size distribution of the matrix. Development is also geared toward tailoring the apparent pore size of matrices for the separation of molecules in specific interesting areas of application (e.g., for polishing of recombinant peptides). Recently, some of these modern media have become available for separation of peptides and proteins of moderate size (Hagel, 1993).

Even though the theoretical framework for the separation mechanism in gel filtration is well understood, there are still some uncertainties as to which molecular size estimate the elution volume reflects (Dubin et al., 1990). It is generally agreed that Stokes radius is a poor measure of molecular size, especially if different types of molecules are being studied. The hydrodynamic volume seems to be a general size parameter for polymers, globular proteins, and DNA of a certain length (Potschka, 1987). It appears that more work needs to be done to resolve this fundamental issue.

Until recently, development of instrumentation for gel filtration has been modest. High-precision pumps and automated systems for injection, fraction collection, and calculation were introduced more than a decade ago. New equipment and technology being introduced includes multiple-angle laser light scattering (MALLS) detectors and in-line viscosity detectors, which provide information about the shape and size of molecules, and electrospray ionization mass spectrometry (ESI-MS), which has the power to provide in-line detection of molecular mass. ESI-MS combined with high-resolution gel-filtration columns provides a very powerful system for characterizing protein dimensions. Molecular size of solutes may also be determined with an in-line light-scattering or viscosity detector combined with a concentration-sensitive detector. It is now possible to obtain the exact molecular mass of small-to-medium-sized proteins by connecting an in-line mass spectrometer to the GF column. The advantages of using a mass-sensitive detector are that the system does not need to be calibrated, fluctuations in flow rate are not detrimental, and solute-matrix interactions or solute-solute interactions will not affect the molecular mass obtained. In this case, the column acts as a high-resolution separating device to resolve the sample components prior to determination of molecular mass by mass spectrometry.

BOULDER COLORADO UNIVERSITY LIBRARY

Critical Parameters

Optimization and scaleup of desalting

Optimization of desalting starts with selection of a matrix of suitable exclusion limit with maximum pore volume and proper particle size. The bed volume is dictated by the sample volume to be purified. The actual elution procedure to be used depends upon the dispersion of the chromatography system, which in turn is related to the particle size and the column length. The particle size will influence the pressure drop over the column; this may be a limiting factor determining the maximum possible flow rate, the maximum column length, and the possibility of scaling up to columns of large diameters (i.e., the pressure specification is often reduced for columns of large diameter). However, in this case the column length can be traded off for an increase in column diameter to keep the bed volume high (also see Strategic Planning, Selecting Matrix and Column for Desalting).

To maximize the sample volume that can be applied to a column

A gel matrix should be selected from which the protein is excluded (see Strategic Planning) and in which the contaminants are eluted close to the total liquid volume (See Basic Protocol 1). It should also have a small matrix volume.

The maximum sample volume that can be applied to a column is directly related to the pore volume of the packed bed (V_i). The pore volume is proportional to the geometric column volume. As a rule of thumb, the maximum sample volume that may be applied for desalting is 25% to 40% of the bed volume for the materials listed in Table 8.3.1. The lower figure applies to small columns (e.g., <10-ml bed volume) and the higher figure to large columns (e.g., 100-ml bed volume). After the bed length is determined on basis of the acceptable pressure drop (which is proportional to the bed length), the sample capacity is maximized by increasing the diameter of the column. This is done because the applicable volume is directly proportional to the cross-sectional area of the column. Limitations in applicable column length may be eliminated by using columns in series ("stacked columns").

The maximum volume that theoretically can be applied for desalting is equal to the separating volume of the system, which is calculated by subtracting the void volume from the total volume (see Basic Protocol 1 for steps to determine these quantities). To account for extra zone broadening effects, it is generally advisable to apply not more than 80% of the separating volume or even less—e.g., 50%—for small columns in which eddy dispersion (see discussion below) will contribute proportionally more to the zone broadening.

To optimize the yield of desalting

Sampling of the protein fraction must be adjusted according to the zone broadening of the sample (i.e., the width of the protein zone). For a well-packed column with a plate height (see Basic Protocol 2) roughly equal to two particle diameters, this parameter may be estimated as the sample volume plus the dispersion of the zone that occurs through different pathways in the packed bed—i.e., eddy dispersion. The peak width (w_b) is described by Equation 8.3.1, where V_t is the total liquid volume of the column, d_p is the particle size of the matrix, and L is the length of the packed bed. Thus, for a 5-ml NAP-5 column (0.9 × 2.8–cm bed dimensions) packed with Sephadex G-25 medium-particle matrix, which has a particle size of 175 µm, eddy dispersion adds ~0.6 ml to the sample volume. This is roughly equal to the maximum sample volume (i.e., the separation volume; see Basic Protocol 1) and thus the eluted fraction is twice the sample volume. However, for a much larger 4 × 85–cm column packed with Sephadex G-25 coarse-particle matrix, which has an estimated particle size of 300 µm, the eddy dispersion adds ~100 ml to the originally applied sample volume of 400 ml, giving a protein zone of 500 ml (see Fig. 8.3.3). If G-25 medium-particle matrix were used, the smaller particle size would in this case theoretically decrease the protein zone from 500 to 475 ml, at the expense of a pressure drop over the column three times higher than with the coarser medium.

$$w_b = 4(V_t)\left(\frac{2d_p}{L}\right)^{\frac{1}{2}}$$

Equation 8.3.1

To minimize the time of the separation

Use the highest possible flow rate for the desalting step. High flow rates are suitable for desalting because eddy dispersion effects are rather insensitive to flow rate. The limiting factors on the flow rate are the rigidity of the matrix and the pressure limitations of the system used. Because media with small pore volume are generally more rigid than those with

RCU IASI CENTRAL UNIVERSITY LIBRARY

large pore volume, the disadvantage of a smaller separating volume may be offset by increasing the applicable flow rate to keep productivity high when working with the former type of medium. Incomplete removal of fine particles from the medium prior to packing may reduce the flow rate of the column after use as a result of partial blockage of the outlet filter. Operating the column with the flow direction reversed may help to increase the applicable flow rate in such situations.

The total process time for repetitive runs to desalt large volumes may be further minimized by utilizing the dead time of the void fraction. To do this, samples are added, not after total liquid volume + sample volume + eddy dispersion volume, but after subtracting the void volume from the total process volume (i.e., the next sample is added before the salt peak of the previous sample has emerged). The saving in time may be in the range of 20% to 30%, but this practice can only be recommended for well-understood processes.

$$R_s = \frac{1}{4}\log_{10}\left(\frac{M_{r1}}{M_{r2}}\right) b \left[\frac{1}{\left(\frac{V_o}{V_i + K_D}\right)}\right] \left(\frac{L}{H}\right)^{\frac{1}{2}}$$

Equation 8.3.2

Optimization and scaleup of protein fractionation

There are two basic ways to optimize the resolution of a separation. One is to increase the peak-to-peak distance and the other is to decrease the peak widths. The resolution for gel filtration is described by Equation 8.3.2, where M_{r1} and M_{r2} are the molecular masses of the two sample components being separated, b is the slope of the selectivity curve ($dV_r/d\log M_r$), V_o is the void volume of the column, V_i is the pore volume of the bed, L the bed height and H the plate height of the column (i.e., the term $L/H = N$, the plate number). The impact of these parameters will be described below.

To increase the peak-to-peak distance

1. A matrix of high selectivity (see *UNIT 8.1*) should be chosen. The most important factor for optimizing protein fractionation is selectivity, which is the inherent ability of the matrix to resolve proteins of similar size. This factor is a function of the pore-size distribution of the matrix. Selectivity is often illustrated by a plot of the elution volumes of a series of reference substances versus the logarithm of the molecular size or molecular mass of those substances. A function of the elution volume (e.g., the distribution coefficient; see Basic Protocol 3) may be substituted for the elution volume in constructing a selectivity curve. The steeper the selectivity curve (i.e., the greater the absolute value of the slope b, defined above), the better the resolvability of proteins differing in size. The selectivity of different gels can be judged from the fractionation range stated by the manufacturer. Gels with a small fractionation range (in terms of $\log M_r$) will generally have higher selectivity than gels with a large fractionation range.

2. A gel of high pore volume should be selected. The pore volume of a matrix is often related to the particle size. Smaller particles need to resist a higher friction force from the high flow rates that are used with matrices of smaller bead size; therefore the matrix volume is generally larger and pore volume smaller when using these matrices. A traditional soft-gel matrix may have up to 98% pore volume whereas rigid matrices such as silica may have as low a pore volume as 52% (of the bead volume).

3. A column of sufficient length (L) should be selected. The resolution between peaks is proportional to the square root of the column length. The column length chosen for protein fractionation is usually optimized with respect to the particle size of the matrix—i.e., a 100-cm column is often chosen where the particle size is 100 μm, a 60-cm column where the bead size is 30 μm, and a 30-cm column where the bead size is 10 μm.

To decrease peak width

1. A material that yields a small void volume (V_o) should be selected. Reducing the void volume will decrease the dispersion of the protein band (as will minimization of all extra column dead volumes). Irregularly shaped or rigid particles generally yield beds with larger void volumes than spherical, nonrigid beads.

2. A material of small particle size should be used as this will affect the plate height (H; also see Basic Protocol 3). Zone broadening of proteins within a column is directly proportional to the square of the particle size. Therefore, smaller particles will yield smaller peak widths unless very large sample volumes are applied or large system effects are causing dispersion of zones.

3. The separation should be run at optimum

BOSTON UNIVERSITY LIBRARY

flow rate, as this will affect the plate height (H; see Basic Protocol 3). Zone broadening is caused by dispersion within the void volume of the chromatographic bed, or by nonequilibrium effects that occur in the separation process itself when the mobile-phase velocity in the extraparticle space is too high as compared with the diffusivities of the proteins and the diffusion distances in the matrix. Flow rates that will allow minimum zone broadening and complete equilibrium will in most cases be impractically low. Table 8.3.7 lists practical flow rates for different solute molecular weights and matrix particle sizes, which represent compromises with regard to these parameters.

Dissolving the sample

Care must be taken to assure that the sample is completely dissolved in the GF buffer. Conditions must be chosen so that spontaneously formed aggregates are dissolved and that aggregates that are to be removed are not dissolved merely to reaggregate in the fraction tube. Because gel filtration permits rather free choice of pH, ionic strength, and additives, it is generally not difficult to find suitable conditions.

Solute-matrix interactions

Solute-matrix interactions are generally to be avoided. Even though mixed-mode separations (i.e., separations on the basis of both molecular size and the solute-matrix interaction) have sometimes given good resolution, one must be concerned about the reproducibility of such mechanisms. Solute-matrix interactions may be specific or nonspecific, and the chemical basis may be an ionic or hydrophobic interaction. In the case of specific solute-matrix interactions, the only thing to do is to change the matrix to one based on a different chemistry. Ionic interaction can be suppressed by addition of an electrolyte; hydrophobic interaction can be reduced by adding an organic modifier (e.g., methanol, acetonitrile, or isopropanol) or by reducing the ionic strength of the buffer. Of course, avoidance of solute-matrix interactions is a prerequisite for estimating molecular size using a calibrated column (see Basic Protocol 3 and Support Protocol). Solute-matrix interactions can be uncovered by varying the composition of the mobile phase or temperature and noting the effect. Because ideal gel filtration is driven by entropy, temperature should not affect the elution volume of proteins (unless the temperature change affects protein size, which is a different issue).

Column maintenance

Column life is drastically reduced by introduction of air into the column, deposition of material on the bed surface, and use of flow rates that exceed the maximum flow rate. To protect valuable columns from air it is advisable to use an air trap, to degas the buffer, or to use a deaerator. Deposition of material is minimized by filtering both sample and buffer through a 0.22-μm filter prior to application. Exceeding the maximum flow rate or pressure (e.g., by failing to compensate for viscosity effects at low temperatures) will result in the bed being compressed. This may be avoided by using a flow or pressure sensor to shut down the pump if the limit is exceeded. To maintain the calibration of a column for molecular size determinations (see Basic Protocol 3 and Support Protocol), the flow through the column should never be turned off completely, but may be reduced (e.g., to 1/10) and the effluent recycled to the buffer reservoir through a filter. The buffer should be treated with preservatives and must be discarded when the column is put back in service. According to the author's experience, the lifetime of a column may be expected to exceed several years and calibration may remain valid for up to 1 year if the appropriate precautions are taken.

A parameter that reflects the possibility of bed compression is the void fraction—i.e., the void volume divided by the bed volume. This parameter is 0.26 for beds made up of tetrahedral, closely packed spheres and 0.40 for beds made up of randomly packed spheres. Experimentally determined void fractions vary from 0.30 to 0.47 because of the surface and packing properties of the matrix and the particle shape. Void fractions <0.30 may indicate compression of beads.

Flow rate

Flow rate must be carefully controlled, especially if the column is used for characterization of molecular size as in Basic Protocol 3. The actual flow rate of a pump seldom agrees with the nominal flow rate, so it is essential that the flow rate be checked regularly.

Troubleshooting

Gel filtration is a comparatively uncomplicated separation technique and troubleshooting is straightforward. Provided that the critical aspects of the technique (see Critical Parame-

BCU IASI - BIBLIOTECA UNIVERSITARĂ

ters) are kept under control, gel filtration can be expected to be relatively trouble-free. Problems that may arise most frequently are compression of the bed and consequent decrease in flow rate, air bubbles in the column, tailing or leading peaks, and peaks that elute in the "wrong place."

Bed compression and decreased flow rate

Compression of the bed most likely results from the maximum flow rate being exceeded during packing (see Critical Parameters). It may also arise from excessive packing of the column, which creates an inhomogeneously packed bed with a denser layer in the lower part of the column that will eventually collapse even at a moderate flow rate. A column found to have a compressed bed must be repacked.

Compression may also result from partial clogging of the inlet filter and the upper part of the bed. To save the column in such a case, the adaptor should be removed, the filter cleaned, contaminated upper matrix removed, the upper part of the bed swirled, and the bed restabilized (see Basic Protocol 1, step 11). If the column still does not perform as expected, it must be repacked.

Decrease in the flow rate often accompanies compression of the bed. If the gel contains fines, these may block parts of the outlet net, causing a decrease in flow rate. Reversing the direction of flow may restore normal conditions. If this is not successful, the column should be repacked after exhaustive decantation of fines (see Basic Protocol 1, steps 3a and 2b).

Air bubbles in column

Small air bubbles in the column may be due to failure to deaerate the buffer or to the introduction of bubbles during sample injection. Moving the column from a cold room directly to a warmer laboratory will also release air bubbles in the column and should not be done. An attempt may be made to remove the bubbles by pumping deaerated buffer through the column in the reverse direction. Heating the buffer slightly is sometimes successful, as the temperatures of the buffer and gel will equilibrate, allowing the warmer buffer to dissolve some air. This may have to be done at least overnight and perhaps for several days to be fully successful. A large air pocket situated right under the adaptor that has not been taken up by the buffer may be removed using the procedures for readjusting the adaptor to the bed surface (see Basic Protocol 1, step 11).

Tailing and leading peaks

Tailing peaks indicate a large dead volume somewhere in the system, solute-matrix interactions, or a bed that is packed too loosely. To check the system for extra dead volumes, another column or a dummy column of small volume should be packed and a sample run to see if the problem reproduces itself. If it does not, the problem can be attributed to the column, and a sample of a different substance should then be run to ensure that the effect is not caused by solute-matrix interactions. If the tailing probem is still noted, the defect probably lies in the column packing, and an attempt should be made to stabilize the bed at a higher flow rate (see Basic Protocol 1, step 11). If this does not help, the column should be repacked at a higher flow rate. Before doing this, however, it should first be ascertained that the tailing is really seen with the sample to be run under the actual conditions under which it is to be run. Slight tailing of a small solute may be negligible when compared to a large diffusion-regulated zone broadening of a protein. Tailing may also be caused by partial separation of a small amount of a solute of slightly smaller size or slightly different surface properties.

Leading (fronting) peaks indicate solute-matrix interactions or cracks in the bed. Solute-matrix interactions (e.g., ionic-exclusion effects) are detected by running a different substance and/or using a different buffer. If the problem is related to the column packing (i.e., cracks in the bed), the column must be repacked at a lower flow rate. As with tailing peaks, however, the actual sample should first be run under the actual conditions to be used to determine whether the effect of fronting on the separation is serious. Fronting may also be caused by partial separation of a small amount of slightly larger contaminant.

Peaks that elute in the wrong place

Peaks that elute prior to the void volume indicate ionic-exclusion effects. To suppress these effects, the ionic strength of the GF buffer should be increased. Solutes that elute after the total liquid volume indicate adsorptive effects. These may be masked by adding an organic modifier or by increasing or decreasing the ionic strength of the buffer (see Critical Parameters for discussion of solute-matrix effects).

Anticipated Results

The results from gel filtration are in general very predictable. Guidelines for optimization

PROPERTY OF CONTRA COSTA COUNTY LIBRARY

and scaleup are presented in Hagel et al. (1989) and Hagel and Janson (1992).

Time Considerations

Gel filtration using media of small particle size may generally be carried out in half an hour. Gel filtration using traditional larger-particle matrices are typically run for 5 hr.

Preparation (i.e., swelling) of the gel will take 3 hr to overnight, depending on the procedure chosen. Packing of the column takes 1 to 4 hr, and the subsequent stabilization of the packed bed takes 1 hr. Thus, a column is typically prepared in a day, and testing of the packing quality is carried out overnight.

Literature Cited

ASTM (American Society for Testing and Materials). 1978. Annual Book of ASTM Standards. ASTM D 3016-78. American Society for Testing and Materials, Philadelphia.

Andrews, P. 1970. Estimation of molecular size and molecular weights of biological compounds by gel filtration. *In* Methods of Biochemical Analysis, Vol. 18 (D. Glick, ed.) pp. 1-53. Wiley-Liss, New York.

Casassa, E.F. 1967. Equilibrium distribution of flexible polymer chains between a macroscopic solution phase and small voids. *J. Polym. Sci. Polymer Letters* Part B 5:773-778.

de Haen, C. 1987. Molecular weight standards for calibration of gel filtration and sodium dodecyl sulfate–polyacrylamide gel electrophoresis: Ferritin and apoferritin. *Anal. Biochem.* 166:235-245.

Dubin, P.L., Kaplan, J.I., Tian, B.-S., and Metha, M. 1990. Size-exclusion chromatography dimensions for rod-like macromolecules. *J. Chromatogr.* 515:37-42.

Flodin, P. 1961. Methodological aspects of gel filtration, with special reference to desalting operations. *J. Chromatogr.* 5:103:115.

Granath, K.A. and Flodin, P. 1961. Fractionation of dextran by the gel filtration method. *Makromol. Chem.* 48:160-171.

Hagel, L. 1985. Effect of sample volume on peak width in high-performance gel filtration chromatography. *J. Chromatogr.* 324:422-427.

Hagel, L. 1989. Gel filtration chromatography. *In* Protein Purification: Principles, High Resolution Methods, and Applications (J.-C. Janson and L. Rydén, eds.) pp. 63-106. VCH Publishers, New York.

Hagel, L. 1992. Peak capacity of columns for size-exclusion chromatography. *J. Chromatogr.* 591:47-54.

Hagel, L. 1993. Size-exclusion chromatography in an analytical perspective. *J. Chromatogr.* 648:19-25.

Hagel, L. and Andersson, T. 1991. A simple procedure for calibrating a column in solute radius. Presented at the 11th International Symposium on HPLC of Proteins, Peptides and Polynucleotides, October 20-23, Washington D.C.

Hagel, L. and Janson, J.-C. 1992. Size-exclusion chromatography. *In* Chromatography, 5th ed., part A: fundamentals and techniques (E. Heftmann, ed.) pp. A267-A307. Elsevier/North-Holland, Amsterdam.

Hagel, L., Lundström, H., Andersson, T., and Lindblom, H. 1989. Properties, in theory and practice, of novel gel filtration media for standard liquid chromatography. *J. Chromatogr.* 476:329-344.

Lathe, G.H. and Ruthven, C.R.J. 1956. The separation of substances and estimation of their relative molecular size by the use of columns of starch in water. *Biochem. J.* 62:665-674.

Lindqvist, B. and Storgårds, T. 1955. Molecular-sieving properties of starch. *Nature* 175:511-512.

Moore, J.C. 1964. Gel permeation chromatography. I. A new method for molecular weight distribution of high polymers. *J. Polym. Sci.* Part A 2:835-843.

Pharmacia Biotech. 1991. Gel Filtration Principles and Methods (5th ed.), Lund, Sweden.

Potschka, M. 1987. Universal calibration of gel permeation chromatography and determination of molecular shape in solution. *Anal. Biochem.* 162:47-64.

Porath, J. and Flodin, P. 1959. Gel filtration: A method for desalting and group separation. *Nature* 183:1657-1659.

Key References

Yau, W.W., Kirkland, J.J., and Bly, D.D. 1979. Modern size-exclusion liquid chromatography, practice of gel permeation and gel filtration chromatography. John Wiley & Sons, New York.

Comprehensive overview of theory and use of gel filtration, covering the entire spectrum of applications (of which protein separation is only a small part). Describes in detail analytical determination of molecular size and size distribution of aqueous and nonaqueous polymers.

Hagel, 1989. See above.

Describes practical implications of gel filtration theory, with special reference to proteins; a good complement to Yau et al. (1979), particularly for protein chemists.

Contributed by Lars Hagel
Pharmacia Biotech AB
Uppsala, Sweden

BCU IASI - CENTRAL UNIVERSITY LIBRARY

Hydrophobic-Interaction Chromatography

There are many separation techniques available to the modern chromatographer, all of which exploit either the physicochemical or biological properties of the target molecule. Chromatographic techniques are generally divided into two broad categories—nonadsorptive and adsorptive—based on whether or not the target molecule interacts with the chromatographic matrix. The major nonadsorptive method is gel-filtration chromatography (*UNIT 8.3*), which separates molecules on the basis of size. Gel filtration is essentially a sieving technique; the molecules being fractionated do not adsorb to or interact with the matrix. All other chromatographic methods—ion-exchange chromatography (*UNIT 8.2*); affinity methods (Chapter 9), including metal-chelate affinity chromatography (MCAC); reversed- and normal-phase chromatography; and hydrophobic-interaction chromatography (HIC; the subject of this unit)—are adsorptive techniques. The basis for separation in adsorptive techniques is the existence of varying degrees of interaction between the chromatographic matrix and the molecules to be resolved.

Hydrophobic-interaction chromatography is used to separate proteins on the basis of the differing strengths of hydrophobic interaction between the proteins and a matrix containing immobilized hydrophobic groups. All proteins are to an extent hydrophobic. Most proteins contain some quantity of the hydrophobic amino acids (i.e., those with nonpolar R groups—alanine, phenylalanine, valine, tryptophan, leucine, isoleucine, and methionine). These amino acids are generally considered to be buried within the protein molecule as a result of the aqueous environment. However, because of folding constraints, some nonpolar amino acids are found on the surface of the molecule as distinct hydrophobic patches. These patches of hydrophobicity are most likely important in the maintenance of protein structure, and thus in determining the biological function of the molecule. The number of such hydrophobic patches and their degree of hydrophobicity is characteristic of a particular protein. The variation in relative hydrophobicity from one protein species to another makes it possible to selectively adsorb and elute proteins to and from a suitable hydrophobic matrix.

Application of HIC involves loading a protein mixture onto a matrix containing hydrophobic ligands in a high-salt environment, then eluting the proteins by decreasing the salt in the environment (e.g., using a descending salt gradient) or by some other method of adjusting the polarity of the water phase (e.g., by adding nonionic detergents or organic solvents).

HIC is often used in the early stages of a separation scheme, especially after a protein precipitation step carried out in the presence of salt, because the binding conditions used in HIC involve high salt concentrations. Elution conditions for HIC involve low salt concentrations; therefore the eluate from an HIC step is ready, in most cases, for application to an ion-exchange column, where binding occurs at low salt concentrations. For this reason, HIC and IEC are often used one after the other in purification strategies. Because binding of the target molecule to an HIC matrix is facilitated by high salt concentration, rather than by organic solvents as in reversed-phase chromatography, HIC is a mild technique that does not destroy the biological activity of the target molecule. HIC, IEC, and gel filtration are useful techniques to consider when developing a purification strategy for labile molecules whose biological activity would be affected by other separation methods.

The Strategic Planning section discusses the important parameters in designing and optimizing a separation by HIC, including preparing the sample and choosing a matrix, column, and buffer. Protocols are provided for packing and testing a column (see Basic

Copyright © 1995 by John Wiley & Sons, Inc.

CU ANSCHUTZ CENTRAL UNIVERSITY LIBRARY

Protocol 1 and Basic Protocol 2 and Alternate Protocol); determining binding and elution conditions (see Support Protocol 1); eluting the sample (see Basic Protocol 3); and cleaning, regenerating, and storing HIC columns (see Support Protocol 2).

STRATEGIC PLANNING

The parameters to consider when selecting a matrix for hydrophobic-interaction chromatography (HIC) and optimizing a separation based upon hydrophobic interactions are: (1) ligand type and degree of substitution of the ligand in the matrix, (2) composition of the matrix support, (3) type and concentration of salt used in the buffer, and (4) pH, temperature, and additives used for the buffer. Table 8.4.1 lists some commercially available media for HIC. The most appropriate matrix will be one that allows the greatest amount of contaminant material to pass through the column during sample loading and initial rinsing. If scale-up is planned for the separation, the cost of items such as buffer additives and waste disposal should be considered for optimum process economy; for economic reasons the most appropriate matrix in such cases will also be one that allows use of the lowest concentration of salt in the binding buffer.

Table 8.4.1 Representative Matrices for Hydrophobic-Interaction Chromatography

Matrix	Matrix support composition	Supplier[a]
45 to 165–μm bead diameter (for capture and intermediate purification)		
Butyl Sepharose 4 Fast Flow	Agarose	PB
Phenyl Sepharose 6 Fast Flow	Agarose	PB
Octyl Sepharose 4 Fast Flow	Agarose	PB
Methyl Macro-Prep	Methacrylate	BR
t-Butyl Macro-Prep	Methacrylate	BR
Bio-Beads SM2, 4, 16, 7	Polystyrene divinylbenzene	BR
Toyopearl butyl	Methacrylate	TH
Toyopearl phenyl	Methacrylate	TH
30 to 50–μm bead diameter (for intermediate and polishing)		
Phenyl Sepharose High Performance	Agarose	PB
Bio-Beads SM	PS/DVB	BR
Toyopearl ether	Methacrylate	TH
Toyopearl butyl	Methacrylate	TH
Toyopearl phenyl	Methacrylate	TH
TSK ether 5-PW	Silica	TH
TSK phenyl 5-PW	Silica	TH
10 to 30–μm bead diameter (for polishing and analytical applications)		
Phenyl Superose	Agarose	PB
Alkyl (neopentyl) Superose	Agarose	PB
Toyopearl ether	Methacrylate	TH
Toyopearl butyl	Methacrylate	TH
Toyopearl phenyl	Methacrylate	TH
TSK ether 5-PW	Silica	TH
TSK phenyl 5-PW	Silica	TH

[a]BR, Bio-Rad; PB, Pharmacia Biotech; TH, TosoHaas. For addresses and phone numbers of suppliers, see *SUPPLIERS APPENDIX*.

[b]See *UNIT 8.1* for definitions of the terms capture, intermediate purification, and polishing.

FCU IASI / CENTRAL UNIVERSITY LIBRARY

If the matrix does not seem to work at all at first, it should be tested again at a later stage in the purification scheme. It is valuable to perform a number of empirical screenings and pilot experiments (see Support Protocol 1) at each step in the purification before establishing a complete multistep separation scheme. Each of these experiments should test several different matrices. See *UNIT 8.1* for discussion of strategies in developing an overall separation scheme.

Ligand Type and Degree of Substitution

The type of immobilized ligand determines the selectivity of the HIC matrix. Generally, two types of ligand are used in HIC—alkyl-chain ligands of various lengths and the aryl (usually phenyl) ligand. Alkyl-chain ligands show true hydrophobic character. At constant ligand-substitution levels, the protein-binding capacities of HIC matrices increase with increasing alkyl-chain length—i.e., with increasing hydrophobicity. Matrices containing the aryl ligand demonstrate a type of mixed aromatic and hydrophobic interaction, because the phenyl ring has the potential for π-π interactions.

The protein-binding capacity of the matrix will increase with increasing degrees of substitution of the immobilized ligand. However, at very high levels of ligand substitution the apparent binding capacity of a matrix will reach a plateau, although the strength of binding will continue to increase in proportion to the degree of substitution. Thus, proteins bound to highly substituted matrices will be difficult to elute. It is worth noting that binding of the surface of a protein to an immobilized HIC ligand is a multipoint attachment, although the attachment between a peptide and an immobilized ligand is a single-point interaction.

Composition of Base Matrix

The support (base matrix) upon which the ligand is immobilized contributes to the selectivity of the HIC medium. The two most common base matrices in use for HIC are 4% and 6% cross-linked agarose, which are strongly hydrophilic carbohydrates, and the synthetic copolymer materials. The selectivity of an agarose-based matrix and a polymer-based matrix substituted with the same ligand will be different. It is possible to achieve the same results using either type of support, but only if the the adsorption and elution conditions are modified appropriately.

It is important to realize when evaluating matrices for HIC that a number of coupling chemistries can be used to attach the alkyl or aryl ligand to the base matrix. For HIC, a matrix should be employed that is prepared with a coupling chemistry in which no charge is imparted to the point of attachment between the base matrix and the hydrophobic ligand. Charge-neutral matrices are often prepared using ether coupling between the base matrix and the hydrophobic ligand. Ligands attached via cyanogen bromide chemistry will be charged, and should be avoided for true HIC.

A third consideration relative to the base matrix is the particle size. The smaller the particle, the higher the resolution achieved in the separation. Matrices with particle sizes ≤34 μm are considered high-resolution media. *UNIT 8.3* contains good discussion and references on the relationship between particle size and resolution. The practical drawback to using high-resolution media for all applications is that the pressure drop across a column bed is inversely proportional to the square of the matrix particle size. Thus, if larger columns (e.g., with bed heights >30 cm) are packed using a 34-μm matrix and the same pump, the back-pressure will increase, the flow rate will decrease, and the time required to achieve the separation will increase. Another complication is the contribution to pressure drop from the viscosity of the sample feed stream and wash buffers used in HIC, which often have a high salt content and are therefore more viscous than water at room

BOURNEMOUTH UNIVERSITY LIBRARY

temperature. Use of a gel matrix with particle size ≥90 µm allows use of reasonably high flow rates at lower back-pressures (e.g., 400 cm/hr at 1 bar). If two HIC steps are used in a separation scheme, it is practical to use a 90-µm matrix early in the purification, and a 34-µm matrix towards the end.

Choosing a Column, Pump, Detector, and Other System Components

Like most adsorptive techniques, HIC is carried out in short columns. A typical HIC column is packed to a bed height between 5 and 15 cm, but the bed height may be as long as 30 cm. Scale-up of an optimized separation is done by increasing column diameter while keeping the bed height constant. Wide columns (with an internal diameter between 1.6 and 5.0 cm) are appropriate for HIC. To achieve the appropriate bed height, it is helpful if the column is equipped with one or two flow adaptors.

A pump should be chosen that is capable of delivering the appropriate flow rate required for HIC, typically between 100 and 600 cm/hr. Because of the high-salt buffers used, peristaltic or syringe pumps are recommended, rather than any type of pump with a ball-check valve. If a pump with a check valve must be used, care must be taken to ensure that no salt crystallizes in the valve seat.

A detector capable of absorbance measurements at 280 nm is used to detect the protein peak in HIC. A chart recorder and fraction collector are recommended.

Type and Concentration of Salt in Buffer

The interaction between the immobilized ligand and the protein is facilitated by the addition of one of the various salting-out salts (i.e., "structure-forming" salts that increase precipitation of hydrophobic compounds; see Fig. 8.4.1) to the equilibration buffer and sample. As the concentration of salt is increased for a particular sample, the amount of protein bound to the immobilized ligand also increases, almost linearly, up to a specific salt concentration. The proteins are then eluted from the column by decreasing the salt concentration, in either a step or continuous descending gradient (see Basic Protocol 3).

The effect of salts in hydrophobic interaction can be accounted for by reference to the well-known Hofmeister (lyophilic-lyophobic) series of anions and cations. Figure 8.4.1 reproduces the Hofmeister series for some anions and cations in terms of the degree to which they promote or inhibit protein precipitation. The most commonly used salts for HIC buffers are Na_2SO_4, NaCl, and $(NH_4)_2SO_4$, all of which have a high salting-out (structure-forming) effect as seen in Figure 8.4.1. Salts can also be listed in order of their effect on the surface tension of water. For more detailed discussion of the Hofmeister

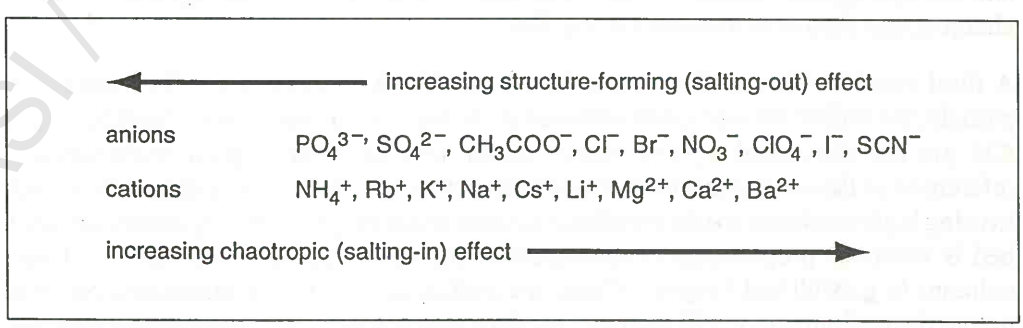

increasing structure-forming (salting-out) effect
anions
PO_4^{3-}, SO_4^{2-}, CH_3COO^-, Cl^-, Br^-, NO_3^-, ClO_4^-, I^-, SCN^-
cations
NH_4^+, Rb^+, K^+, Na^+, Cs^+, Li^+, Mg^{2+}, Ca^{2+}, Ba^{2+}
increasing chaotropic (salting-in) effect

Figure 8.4.1 The Hofmeister series illustrating the effects of some anions and cations in promoting protein precipitation. Ions with a higher salting-out effect promote binding to hydrophobic interaction chromatography matrices, whereas ions with a higher salting-in effect promote elution from those matrices.

ECU IAST / CENTRAL UNIVERSITY LIBRARY

series, see Melander and Horvath (1977); for more detail on surface-tension salts, see Srinivasan and Ruckenstein (1980), Franks (1988), and Hodder et al. (1989).

Effect of Buffer pH

In published accounts of hydrophobic-interaction chromatography, pH is shown to be an important separation parameter. Generally, as the pH is increased, hydrophobic interaction is decreased. It is possible that increased titration of the charged groups that occurs with increased pH results in a protein that is more hydrophilic. A decrease in pH tends to increase hydrophobic interaction. Thus, proteins that do not bind to a hydrophobic-interaction matrix at neutral pH do bind at acidic pH. It is advisable to check the binding of the protein of interest at several pH values in the course of optimizing a HIC separation.

Effect of Temperature

On the basis of the hypothesis that interaction between hydrophobic moieties on the protein surface and the immobilized hydrophobic ligand is entropy-driven (see Background Information), one would expect that the interaction should increase with increasing temperature. Experimental evidence to this effect has been presented (Hjerten, 1977), but the role of temperature in HIC is evidently complex, because the opposite results have also been observed (Visser and Strating, 1975). The practical importance of this is that one must be aware that a process developed at room temperature may not be reproducible in the cold room (and vice versa).

Elution of Proteins

Elution of compounds bound to an immobilized hydrophobic ligand can be achieved in one of three ways: (1) by a linear or stepwise decrease in the concentration of salt, (2) by addition of some proportion of organic solvent to the elution buffer, or (3) by addition of neutral detergents to the elution buffer.

The most common way to desorb the proteins from the HIC support is to reduce the salt concentration of the buffer (see discussion above under Type and Concentration of Salt in Buffer), or to use water alone as an eluant. An alternative elution scheme for desorption of the bound protein may exploit the fact that low concentrations of water-miscible alcohols, detergents, and salting-in salts (i.e., chaotropic salts that decrease precipitation of hydrophobic compounds; see Fig. 8.4.1) result in the weakening of the protein-ligand interaction. The nonpolar regions of the alcohols or detergents compete with the bound protein for the hydrophobic ligand, resulting in displacement of the protein. Salting-in salts disrupt the structure of the water, decrease the surface tension, and thereby weaken the hydrophobic interaction. These compounds can be used in the elution scheme to affect selectivity during desorption, but there is a risk that the protein will be denatured or inactivated after exposure to them. Detergents can be effective for cleaning HIC media that have very strongly hydrophobic proteins bound to them (see Support Protocol 2).

Preparation of High-Salt Buffers

As discussed above under Type and Concentration of Salt in Buffer, the binding of a protein to an HIC medium is facilitated by high salt concentrations. The sample should thus be in high salt prior to application to the HIC column. This can be done by direct dilution in high-salt solution or by buffer exchange through gel filtration (see *UNIT 8.3*). Before deciding to use gel filtration, it is necessary to ensure that none of the sample components precipitate in the high-salt buffer. This can be done in a small test tube by adding an aliquot of the high-salt buffer to the sample and checking for flocculation.

BRANDEIS UNIVERSITY LIBRARY

As the use of ammonium sulfate in buffers for HIC is very common, preparation of such buffers requires some comment. Because ammonium sulfate can be contaminated with heavy metals, iron in particular, the best grade available (usually called enzyme grade) should be used. For the same reason it may be necessary to include micromolar amounts of EDTA in HIC buffers containing ammonium sulfate. Preparation of a saturated solution takes several hours; therefore it is best to prepare the solution a day before chromatography is to be attempted. Concentrations of ammonium sulfate are expressed in terms of molarity or percent saturation. For use in a saturation table, the sample can be considered to be 0% saturated or 0 M salt.

Scopes (1994) presents two equations for calculating ammonium sulfate concentration. Also useful are the salting-out tables found in general biochemistry references including Segel (1976), which is aimed at students, and Englard (1990). Before buffer preparation, it is important to take into account that addition of ammonium sulfate to water causes an increase in volume; e.g., 761 g of ammonium sulfate is added to one liter of water to make a saturated solution, resulting in a final volume of ~1.5 liters. If the water is heated to between 70°C and 80°C, the ammonium sulfate will go into solution faster (in 30 min to 1 hr). If the solution is heated, it should be cooled before use—e.g., overnight in a refrigerator. In a saturated solution there will be crystals of ammonium sulfate in the bottom of the flask. The pH of a saturated ammonium sulfate solution is ~5; it can be adjusted with concentrated ammonium hydroxide. The aliquot of solution to be used should first be filtered through filter paper. Saturated ammonium sulfate solution may be stored several weeks at 4°C.

NOTE: All glassware used in hydrophobic-interaction chromatography should be clean. Spills of highly concentrated salt solutions should be cleaned up immediately, especially if the salt splashes onto instruments.

BASIC PROTOCOL 1

PACKING A COLUMN WITH AN HIC MATRIX OF ≥90-µm BEADS

This protocol describes preparation and degassing of the HIC gel medium and packing of the column. After these procedures are completed, a packed column will be ready for testing (Basic Protocol 2) and sample application. This protocol and that for high-resolution media (see Alternate Protocol) are generally applicable to all bulk carbohydrate-based media, and can be skipped if a prepacked column is to be used. If non-carbohydrate-based media are to be used, consult the manufacturer's literature for packing instructions.

Materials

Appropriate HIC matrix with bead size ≥90 µm (see Strategic Planning and Support Protocol 1)
Packing solution: distilled H_2O or low-ionic-strength buffer (e.g., 20 mM sodium phosphate, pH 7.0)
20% (v/v) ethanol

Chromatography column with flow adaptors (see Strategic Planning)
Sintered-glass funnel (medium or coarse grade)
Carpenter's level
Pump (see Strategic Planning)

1. Equilibrate all materials to the temperature at which chromatography will be performed.

2a. *To wash the HIC matrix by decantation:* Allow gel to settle, decant off the preservative solution, and add an excess of packing solution. Repeat wash several times.

ECU IASI CENTRAL UNIVERSITY LIBRARY

2b. *To wash the HIC matrix by filtration:* Pour the gel into a sintered-glass funnel attached to a suction flask and vacuum line and allow the preservative solution to filter off. Release the suction, resuspend the gel in an excess of packing solution, and restore the suction to filter off the packing solution. Repeat wash several times.

Gels are shipped in a preservative solution—usually 20% ethanol or a dilute thimerosal or sodium azide solution.

Washing by filtration is a faster method than washing by settling and decantation. A sintered-glass filter must be used, not filter paper in a Buchner funnel.

3. Prepare a slurry consisting of 70% washed, settled gel in packing solution.

The packing solution (water or dilute buffer) provides a low-viscosity liquid phase for packing the gel in the column.

4. Degas the gel slurry 3 to 5 min under vacuum with swirling.

A water aspirator pulls enough vacuum for this degassing.

5. Mount column vertically. Flush the end pieces of the column with 20% ethanol, then flush with packing solution to replace the ethanol. Close outlet, leaving a few centimeters of packing solution in the bottom of the column.

6. Pour the gel slurry into the column with one motion, avoiding introduction of air bubbles.

Pouring the slurry down a glass rod held up against the side of the column wall will minimize the introduction of air into the slurry. In the absence of a glass rod, pouring the gel down the column wall will have the same effect.

7. Immediately fill the remainder of the column with packing solution, mount the top adaptor, and connect the column to the pump. Use a carpenter's level to check that the column is vertical.

It is not necessary to push the adaptor down into the gel at this point, but it is convenient to move it into the column ~1 cm above the settled gel to minimize the need for adjustment after the matrix is packed.

8. Fill the buffer reservoir with packing solution. Open the bottom of the column and set the pump to run at the desired flow rate. Continue passing the packing solution through the column until a constant bed height is reached, then pass an additional 3 bed volumes of solution through at the same flow rate.

Different flow rates are recommended by manufacturers for different gel matrices—e.g., Sepharose 6 Fast Flow matrices are packed at a flow rate of 500 cm/hr, and Sepharose 4 Fast Flow gels at a flow rate of 400 cm/hr. Ideally, these two media should also be packed at constant pressure: 1.5 bar for Sepharose 6 Fast Flow and 1.0 bar for Sepharose 4 Fast Flow. However, constant-pressure packing requires a pressure gauge in the chromatography system and this is not always available. If the recommended flow rate cannot be achieved, the gel should be packed at the maximum flow rate that the pump can deliver.

IMPORTANT NOTE: *Do not exceed 70% to 75% of the packing flow rate in subsequent chromatographic procedures, because if the separation is run at a flow rate approaching that used for packing, there is a risk that the column will continue to pack during sample application and elution, leading to distorted peaks and non-reproducible elution patterns.*

9. Mark the level of the packed bed on the column tube. Close the outlet of the column, then stop the pump. Disconnect the inlet tubing from the pump. Adjust the adaptor to the top of the gel medium and lower it 3 mm below the mark on the column tube. Tighten the adaptor and reconnect the inlet tubing to the pump.

The column is now ready to use in a separation. For examples of protein purifications done by HIC, see Background Information.

BOSTON UNIVERSITY LIBRARY

ALTERNATE PROTOCOL

PACKING A COLUMN WITH AN HIC MATRIX OF ≤34-µm BEADS

This protocol describes the variations in packing procedure for high-resolution HIC matrices. Carry out Basic Protocol 1 with the following additional materials and alternate procedures at the indicated steps.

Additional Materials (also see Basic Protocol 1)

Appropriate HIC matrix with bead size ≤34 µm (see Strategic Planning and Support Protocol 1)
Tween 20
Chromatography column rated at 5 bar, with flow adaptors (see Strategic Planning)
Pressure gauge

3a. Add 250 µl (one drop) Tween 20 per 500 ml of gel slurry.

Tween 20 will decrease the surface tension within the gel solution, making the slurry more homogeneous and easier to swirl and degas.

7a. Pack the column at a flow rate of 30 cm/hr until the bed reaches a constant height (or at least 20 min). Shut off the pump, then lower the adaptor to ~1 cm above the settled gel bed. Turn on the pump and increase the flow rate until a pressure of 5.0 bar is reached. Continue packing at this pressure for 30 to 60 min.

After the high-resolution gel is packed, some resistance may be encountered from the packed bed when attempting to push the adaptor down 3 mm into the bed (see Basic Protocol 1, step 9). It is nevertheless important to do this before using the column in a separation.

BASIC PROTOCOL 2

TESTING THE PACKED BED

Testing the packed column is easy, and is recommended for two reasons. First, it is better to find out whether or not the column is properly packed before the sample is applied. It is not unusual for someone who has never packed a column before to try two or three times before succeeding. Second, testing column performance with a tracer before the sample is applied and the column used and regenerated provides a baseline for comparison after a number of runs. Often, the theoretical plate number and/or asymmetry factor, determined as described below, can be used to predict column failure. Thus, it is recommended that the packed column be tested before it is used and then again before each additional use (or before every other use, or on some other regular schedule). To test a packed column, a low-molecular-weight test substance that has no interaction with the matrix is injected into the column. The number of theoretical plates (N) or the height equivalent to a theoretical plate (H, sometimes called HETP) is then calculated from the elution volume of the test substance. The asymmetry factor (A_s) is also calculated. The small molecular weight assures that the test substance has access to the entire chromatographic bed. 1% (v/v) acetone can be used as a test substance. The acetone sample is applied just as though it were any other chromatographic sample. The sample volume used for this test should be small: 2% of the column volume for 90-µm packings or 0.5% of the column volume for 34-µm packings. However, for evaluating column packing it is appropriate to minimize the volume between the injection point and the detector; thus, short lengths of tubing should be used in the chromatography system for the evaluation step even if longer lengths will be used for the chromatography of an unknown sample later on. The flow rates used for evaluation of the column should be kept low to minimize the zone broadening (*UNIT 8.1*) at the front and rear of the zones. For columns packed with 90-µm matrices, the flow rate should be 15 to 30 cm/hr; for columns packed with 34-µm matrices, the flow rate should be 30 to 60 cm/hr.

ECU IAS/CENTRAL UNIVERSITY LIBRARY

Because acetone absorbs at 280 nm, it is not necessary to change the wavelength setting on the detector to test the packed bed. However, it will be necessary to change the sensitivity of the detector to keep the peak on-scale so that the peak apex can be seen. The variables N, H, and A_s should not be calculated from a peak that has gone off-scale.

The following standard chromatographic equations are used to calculate the number of theoretical plates (N; also see discussion of column efficiency in *UNIT 8.1*) and then to calculate plate height (H):

$$N = 5.54(V_e/W_h)^2$$

and finally $H = L/N$

where V_e is the volume eluted from the start of sample application to the peak maximum, W_h is the peak width measured at half peak height (as traced by the recorder), and L is the length of the packed bed.

Because the plate count (N) is dimensionless, the measurement of V_e and W_h can be taken in distance on the chart paper (mm) or volume of eluant (ml). The units used should be the most convenient for the investigator, but both V_e and W_h should be expressed in the same units for a given calculation. Figure 8.4.2 shows a chromatogram with V_e and W_h marked.

As a general rule, a good value for H is about two to three times the mean bead diameter of the gel being packed. Thus, a good H value for a 90-µm matrix is between 0.018 and 0.027 cm and a good H value for the 34-µm matrix is between 0.0070 and 0.010 cm.

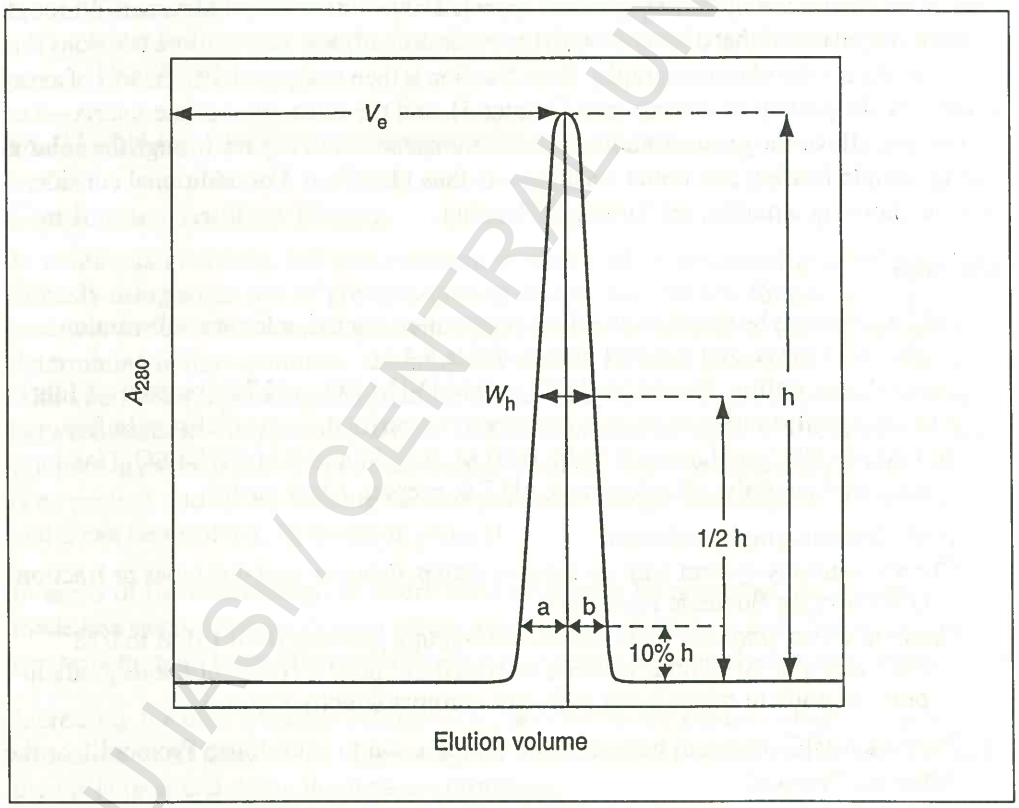


Figure 8.4.2 Chromatogram illustrating parameters used in calculating plate number and asymmetry factor for a column. V_e is the elution volume for the protein whose peak is depicted, h is the height of that peak, W_h is the width of the protein peak at half the peak height, *a* is the width of the leading portion of the peak at 10% of the peak height, and *b* is the width of the tailing portion of the peak at 10% of the peak height.

B.C.U. „M. Eminescu" IASI
CHIMIE
U IASI / CENTRAL UNIVERSITY LIBRARY

Another standard chromatographic data analysis procedure for testing a packed bed is calculation of the asymmetry factor A_s. On the same peak used to measure H on the chart recording, a line is drawn from the peak maximum to the baseline, dividing the peak in two parts lengthwise. Asymmetry is calculated simply by the equation:

$$A_s = b/a$$

where a is the width of the front (leading) portion of the peak measured at 10% peak height and b is the width of the back (tailing) portion of the peak measured at 10% peak height. The chromatogram in Figure 8.4.2 depicts the quantities a and b graphically.

A_s should be as close as possible to 1, but a reasonable value for a short-bed HIC column is between 0.80 and 1.80. An extensive leading edge on a peak ($A_s < 0.80$) is usually an indication the column is packed too tightly, and an extensive trailing edge ($A_s > 1.80$) is an indication the column is packed too loosely. These problems can be solved by repacking the column.

SUPPORT PROTOCOL 1

PILOT EXPERIMENT TO CHOOSE A MATRIX AND DETERMINE BINDING AND ELUTION CONDITIONS FOR HIC

This protocol describes an experiment to identify the appropriate HIC matrix and initial salt concention to bind and elute a target molecule. This experiment is done with very small columns and consumes very little sample. A small quantity of the sample solution dissolved in a high-salt start buffer is applied to several different columns, each packed with a different HIC matrix. Elution is then carried out using a either a descending continuous gradient or a descending step gradient of salt concentration (see Basic Protocol 3 for more discussion of the two gradient types). The results will include a wash-through fraction (i.e, material that did not bind to the column at all) and one or more fractions that will elute during the chromatography. Each fraction is then analyzed by some sort of assay to identify the protein of interest (see Chapter 3), and the most appropriate matrix—i.e., the one that allows the greatest amount of contaminant material to pass through the column during sample loading and initial washing—is thus identified. For additional considerations in choosing a matrix, see Strategic Planning.

Materials

HIC matrices to be tested (e.g., a low-substitution phenyl, a higher-substitution phenyl, a butyl, and an octyl matrix; Table 8.4.1)
General start buffer: 50 mM Na_2HPO_4/1.0 M $(NH_4)_2SO_4$, pH 7.0 (prepare ~1 liter)
5 to 10 mg/ml solution of protein mixture to be purified, in general start buffer
50 mM Na_2HPO_4containing 0.75 M, 0.50 M, 0.25 M and 0 M $(NH_4)_2SO_4$ (for step gradient only; all solutions at pH 7.0; prepare 1 liter each)

1-ml chromatography columns
Chromatography system with syringes or pump, detector, and test tubes or fraction collector (see Strategic Planning)
Gradient mixer (*UNIT 8.2*) set to deliver descending gradient from 1.0 M to 0 M $(NH_4)_2SO_4$ in 10 column volumes of Na_2HPO_4, pH 7.0 (for continuous gradient only; if gradient mixer is not built into chromatography system)

1. Pack each HIC matrix to be tested into a 1-ml column as in Basic Protocol 1 or the Alternate Protocol.

2. Equilibrate each column with at least 10 column volumes of general start buffer.

   The flow rate for equilibration can be fast—e.g., 1 column volume/min. See Critical Parameters for guidelines on column equilibration.

PCU IASI CENTRAL UNIVERSITY LIBRARY

3. Apply 1 ml of protein solution to be purified in start buffer to each column. Perform initial washes in start buffer, and monitor the eluant using a detection scheme appropriate to the protein of interest.

   A slower flow rate is recommended for sample application and elution than for equilibration—e.g., 0.25 column volumes/min.

4. Elute the sample using either a gradient mixer set to deliver a descending gradient from 1.0 M to 0 M $(NH_4)_2SO_4$ in 10 column volumes of Na_2HPO_4, pH 7.0 *or* a step gradient of 10 column volumes each of 0.75 M, 0.50 M, 0.25 M, and 0 M $(NH_4)_2SO_4$ in 50 mM Na_2HPO_4, pH 7.0. Continue to monitor the eluant and construct a chromatogram.

   See Critical Parameters and Troubleshooting for guidelines on interpreting chromatograms obtained from this pilot experiment and for optimizing chromatography conditions.

   The most appropriate matrix is the one that allows the greatest amount of contaminant material to pass through the column during sample application and initial washing. If the protein of interest elutes during sample application or initial washing with start buffer, repeat the procedure with the same gel matrix but increase the salt concentration of the start buffer. Alternatively, more hydrophobic matrices may be investigated. The purification of the protein should be evaluated even where elution occurs in the initial wash, because a considerable amount of contaminating material may be bound to the HIC media and hence removed.

   If the protein of interest does not elute until a 0 M salt concentration has been reached, the salt concentration of the start buffer may be decreased from 1.0 M $(NH_4)_2SO_4$ (e.g., to 0.25 M). Elution may then be done using 50 mM Na_2HPO_4, pH 7.0, or water alone. If the activity of the protein of interest is adversely affected by low salt concentrations, a less hydrophobic matrix that allows elution at a higher salt concentration may be investigated.

ELUTION OF PROTEINS FROM HIC COLUMNS

BASIC PROTOCOL 3

Continuous-Gradient Elution

In continous gradients, salt concentration is increased or decreased gradually and continously using some sort of gradient-mixing device (see *UNIT 8.2*). Simple linear gradients (see *UNIT 8.2* for definitions of gradient types) are the first choice in pilot experiments for determining initial conditions (see Support Protocol 1). After more experience is obtained with a particular separation, the gradient can be made more complex to optimize resolution between specific sample components. This is illustrated in Figure 8.4.3, where the linear gradient in panel A shows very poor resolution of the first two components in the mixture (i.e., peaks 1 and 2). By making the first portion of the gradient slope less steep, peaks 1 and 2 can be resolved, as shown in panel B.

In areas of the chromatogram where resolution must be improved, the gradient can be made less steep, whereas in areas where resolution is adequate or not critical, the gradient can remain more steep. The general effect of gradient slope is illustrated in Figure 8.4.4.

Increasing the total gradient volume—i.e., decreasing the gradient slope—will improve resolution in all parts of the chromatogram. The drawback to this is that it will increase the cycle time and dilute the separated fractions.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

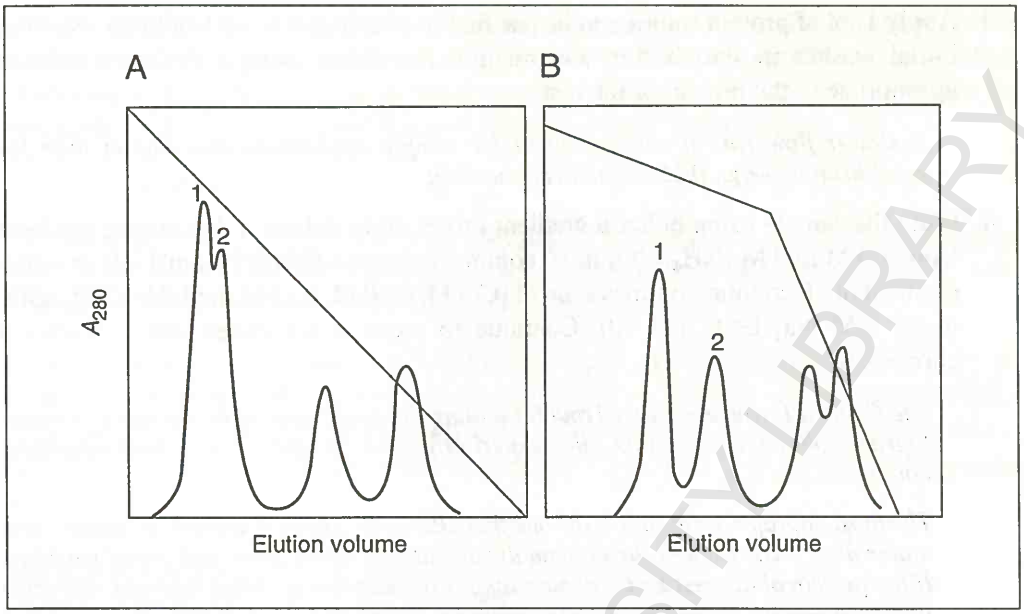


Figure 8.4.3 Chromatograms illustrating use of complex gradient elution to improve resolution. (**A**) A simple linear gradient gives poor resolution of protein peaks 1 and 2. (**B**) A complex gradient where the gradient slope is less steep in the region of peaks 1 and 2 gives better resolution of these two components.

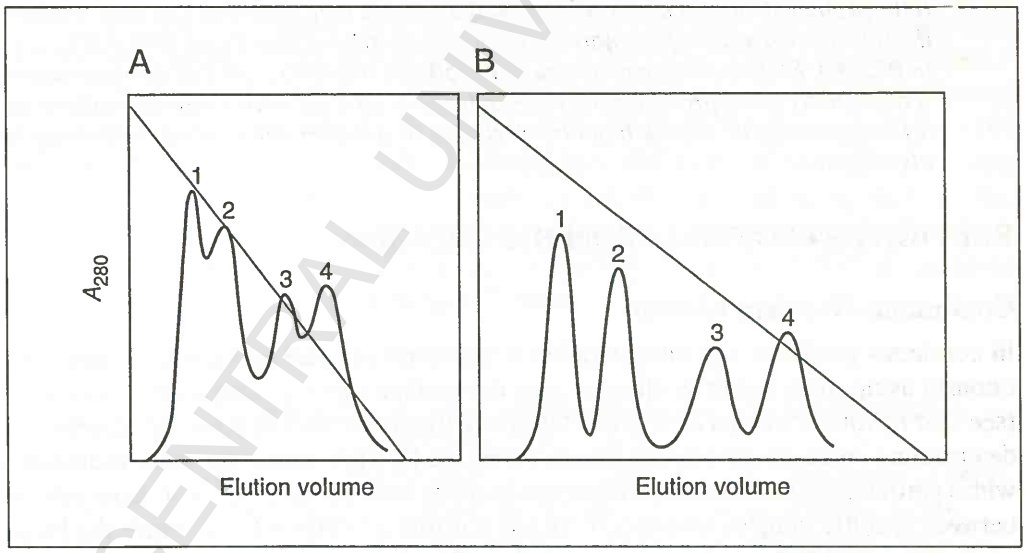


Figure 8.4.4 Chromatograms illustrating the general effect of a change in gradient slope on peak resolution. (**A**) A steep slope gives poor resolution among the four sample components. (**B**) A less steep slope gives better resolution (though the total elution volume is higher, meaning that the eluted sample components will be more dilute and that more buffer is needed for the elution).

Step-Gradient Elution

In step-gradient elution, one or more solutions of a particular discrete concentration are passed through a column. The principle of stepwise elution is to provide maximum resolution in the area of the gradient where the protein of interest elutes. Stepwise elution is preferred for large-scale preparative applications because it is technically simpler and more reproducible than continous-gradient elution. It is also advantageous in small-scale applications because the compound of interest can be eluted in a more concentrated form, provided there is no coelution of a more strongly bound compound. When stepwise elution is used, one must watch out for artifact peaks that can appear if the next step is started too

PCU IASI CENTRAL UNIVERSITY LIBRARY

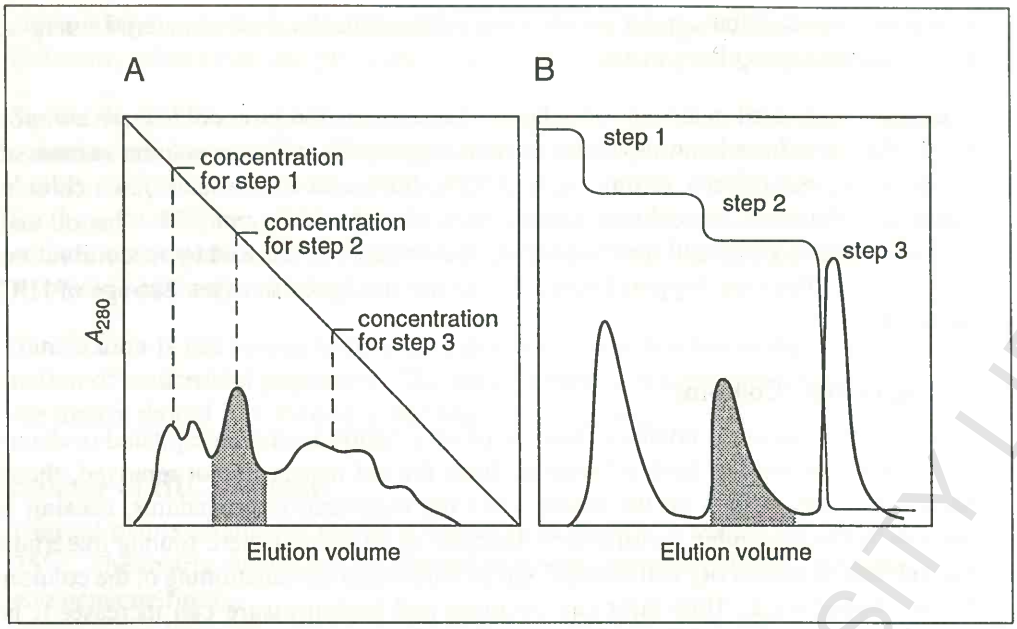


Figure 8.4.5 Chromatograms illustrating use of the results of linear-gradient elution to devise a step-gradient elution scheme. The shaded area represents the target protein peak. (**A**) Elution by linear gradient yields poor resolution among sample components. However, it is possible to note the salt concentration at the points on the gradient where the peaks for the loosely bound contaminants, target protein, and tightly bound contaminants emerge. (**B**) Step-gradient elution with three discrete solutions prepared on the basis of the elution salt concentrations noted in panel A. The target protein peak and the two contaminant peaks are sharply resolved.

early—i.e., during a trailing peak. It is recommended that a stepwise elution scheme be developed after several continuous gradient schemes have been run to characterize sample behavior on the matrix. This development should include three stages. In the first stage a continuous gradient is optimized to elute all the compounds that bind to the gel less tightly than the protein of interest. The elution strength and gradient volume should be high enough to elute all weakly binding contaminants but not high enough to elute the protein of interest. In the second stage the elution strength is increased so that the protein of interest elutes with minimal dilution, but more tightly bound contaminants are not eluting. In the third step the elution strength is increased to elute the most tightly bound contaminants. Once these three regions of the chromatogram are identified, these gradient positions are converted to steps (i.e., solutions representing the concentration of eluant at these points in the chromatogram are prepared and used for step elution). Figure 8.4.5 illustrates this process of switching from a continuous to a step gradient.

REGENERATION, CLEANING, AND STORAGE OF HIC COLUMNS

SUPPORT PROTOCOL 2

Regeneration of HIC Columns

After each use, the matrix must be regenerated to restore its original functionality. There are two general regeneration procedures for HIC columns. The first is to simply wash the column with several column volumes of distilled water (Milli-Q purified or equivalent). The second is to wash the columns with at least two column volumes of 6 M urea or 6 M guanidine hydrochloride and then at least 10 column volumes of start buffer (see Support Protocol 1) or storage solution (see Storage of HIC columns, below). The first procedure is sufficient if the buildup of tightly bound protein contaminants is minimal; the second procedure is recommended for more tightly bound contaminants. Deciding which to use

BOWLING GREEN STATE UNIVERSITY LIBRARY

is situation-dependent, but a quick screen using each, noting if a peak of material emerges, can indicate the appropriate choice.

If detergents were used in the elution scheme, the regeneration protocol is more complicated. In this case the column should be washed sequentially with one column volume of distilled water; one column volume each of 25%, 50%, and 95% ethanol; two column volumes of *n*-butanol; one column volume each of 95%, 50%, and 25% ethanol, and finally one column volume of distilled water. This should be followed by reequilibration with the start buffer (see Support Protocol 1) or the storage buffer (see Storage of HIC columns, below).

Cleaning of HIC Columns

Cleaning of HIC columns involves removal of very tightly bound precipitated or denatured proteins, as well as lipid substances, from the gel matrix. If not removed, these contaminants may remain on the matrix after the regeneration procedures, causing a decline in chromatographic performance. Because of buildup of these fouling materials on the gel, matrix selectivity will change, and in some cases the functioning of the column will also change—e.g., flow rates can decrease and back-pressure can increase. It is preferable to clean the matrix while it is still in the column, a practice referred to as "cleaning in place," so that the column does not have to be unpacked, repacked, and tested again. There is no single recommended protocol for cleaning in place; the procedure must be designed with some knowledge of the contaminants present in the feed stream. In general, however, NaOH is a very effective cleaning agent that can be used for solubilizing both denatured and precipitated protein and lipids.

Removal of precipitated proteins

To remove precipitated proteins, a general recommendation is to wash the column with four column volumes of 0.5 to 1.0 N NaOH at a slow flow rate (e.g., 40 cm/hr), then with two to three column volumes of water. After the water wash, the column can be equilibrated with the start buffer (see Support Protocol 1) or put into the storage buffer (see Storage of HIC Columns, below).

Removal of tightly bound proteins and lipids

There are two general recommendations for the removal of strongly bound hydrophobic proteins, lipids, and lipoproteins. The first is to wash the column with four to ten column volumes of ethanol, starting with 30% ethanol and working up to 70%, then with six to ten column volumes of distilled water (Milli-Q purified or equivalent; this should reduce the alcohol content of the eluate to between 10 and 1 ppm). The 30% to 70% ethanol may be replaced with 30% isopropanol. The second procedure is to wash the gel matrix with one to two column volumes of a 0.5% solution of nonionic detergent (e.g., Triton X-100) in 1 M acetic acid, followed by five column volumes of 70% ethanol to remove the detergent. The ethanol is then washed out with six to ten column volumes of distilled water. After the water wash, the column can be equilibrated with the start buffer or put into the storage buffer.

Sanitization of HIC Columns

Sanitization is the inactivation of microbial populations in the gel matrix. NaOH is very effective for sanitizing the matrix in addition to functioning as a general cleaning reagent. A general sanitizing procedure for HIC columns is to expose the matrix to 0.5 to 1.0 N NaOH for 30 to 60 min. The column should then be washed with at least ten column volumes of start buffer (see Support Protocol 1) if it is to be used right away. The large volume of wash solution is required because a large pH adjustment is being made. If the

PCU IASI CENTRAL UNIVERSITY LIBRARY

column is going to be stored it should be washed with storage buffer (see Storage of HIC Columns, below) until the pH of the effluent equals the original pH of the storage buffer.

Sodium hydroxide is widely accepted as an agent for sanitizing chromatographic media, columns, and systems. For this purpose it is used at higher concentrations (e.g., 0.5 N to 1.0 N). At lower concentrations (e.g., 0.01 N) it is an effective bacteriostatic agent, and for the most common gram-negative contamination a good bactericidal effect is also seen at these concentrations. NaOH also destroys endotoxin on the column and is nontoxic as a contaminant in the end product.

Sanitization is not synonymous with sterilization, which refers to the destructive elimination of a microbial population. The manufacturer's recommendations for autoclaving the matrix should be followed if sterilization is required.

Storage of HIC Columns

Unused matrix should be stored in closed containers at a temperature between 4° and 25°C. The matrix should not be allowed to freeze as this will disrupt the bead structure and generate fines.

Used matrix should be stored in the presence of a suitable bacteriostatic agent—e.g., 0.01 N NaOH or 20% ethanol—at a temperature between 4° and 8°C. Used matrix should also be protected against freezing. Packed columns should be stored under the same conditions. For long-term storage, the matrix should be cleaned (see Cleaning of HIC Columns) before equilibration with the storage buffer. Recycling the storage solution or flushing the column with fresh storage solution at least once a week during long-term storage is an effective method for preventing bacterial growth.

Chromatography matrices from Pharmacia Biotech are supplied in 20% ethanol, and this solution can be used as an alternative to NaOH for storing columns. Other manufacturers ship media in other bacteriostatic solutions, which can also be used as alternative storage buffers.

Other bacteriostatic compounds can be added to the storage buffer, including 0.002% chlorhexidine (which can, however, cause precipitation of sulfate ions if mixed with a high-molarity salt such as that used in start buffers for HIC); 0.5% trichlorobutanol (effective in weakly acidic solutions only), 0.005% thimerosal (effective in weakly acidic solutions); 0.001% to 0.01% phenyl mercuric acetate, borate, or nitrate (effective in weakly alkaline solutions); and 0.02% to 0.05% sodium azide.

COMMENTARY

Background Information

Tiselius (1948) laid down the foundations of what today is called hydrophobic-interaction chromatography (HIC). Since then, a number of theories have been proposed to explain this type of chromatography, but no single one has gained universal acceptance. All of the explanations refer to the interactions of the structure-forming salts (see Strategic Planning, Type and Concentration of Salt in Buffer) and the effects that they exert on other solutes and solvents in a system. Porath (1986) posed a general explanation for HIC based upon salt-promoted adsorption. In an extension of earlier work in which the observations of Tiselius were extended and the behavior of hydrophobic molecules in solution were linked to thermodynamic theory (Porath, 1973), they concluded that the driving force of HIC is the gain in entropy arising from the structural changes in the water surrounding the interacting hydrophobic groups. Hjerten (1977) expressed this idea in terms of the thermodynamic relationship $\Delta G = \Delta H - T\Delta S$. The displacement of water molecules surrounding the hydrophobic ligands on the matrix and the hydrophobic amino acids on the surface of the protein leads to an increase in entropy (ΔS), because more water is returned

BOURNEMOUTH UNIVERSITY LIBRARY

to the bulk solution after a hydrophobic interaction occurs. This leads to a negative value for the ΔG of the system, which implies that hydrophobic interaction is thermodynamically favorable. This also implies that the entropy of water is the driving mechanism for hydrophobic interaction; thus, HIC is unique among the adsorptive separations in that it occurs as a result of its environment.

A second explanation for hydrophobic interaction was proposed by von der Haar (1976) and by Melander and Horvath (1977). According to these authors, hydrophobic interaction results from the increase in surface tension of water arising from dissolved structure-forming salts. A third explanation for hydrophobic interactions was proposed by Srinivasan and Ruckenstein (1980). This hypothesis proposes that hydrophobic interaction is a result of the increase in van der Waals attraction between the protein and the immobilized ligand that results from the increase in the ordered structure of water in the presence of structure-forming salts.

Elution by salt gradient

An example of elution by salt gradient was published by Lau et al. (1987) as part of the purification of acid phosphatase from bovine bones by HIC. The active fractions from previous steps were pooled in 100 mM sodium acetate buffer, pH 6.5, and adjusted to 30% saturation by adding solid ammonium sulfate at 4°C with gentle stirring for ≥2 hr. This enzyme solution was applied to a phenyl HIC column equilibrated with 30% ammonium sulfate in 100 mM sodium acetate buffer, pH 6.5. The chromatography was done at room temperature. The column was eluted with a descending salt gradient starting at 30% ammonium sulfate/100 mM sodium acetate, pH 6.5, and ending with 0% ammonium sulfate/100 mM sodium acetate, pH 6.5. The yield was 92% and the purification was 13-fold. This was the last step in a four-step purification protocol and the purity of the isolated enzyme was determined by SDS-PAGE.

HIC with gradient elution offers an alternative in monoclonal antibody purification, provided that the monoclonal antibody is stable in high salt. As most monoclonal antibody purifications start with an ammonium sulfate precipitation step, the antibody will normally be present either in a high salt–soluble fraction or as a precipitate that can be redissolved and applied to a HIC column. Monoclonal antibody samples containing albumin and transferrin are particularly suitable for HIC because the albumin and transferrin can be eluted as sharp peaks, minimizing contamination of the antibody. This was demonstrated by Sparrman et al. (1986), who used an alkyl Superose matrix. A clarified sample was equilibrated with 0.7 M ammonium sulfate by buffer exchange using gel-filtration chromatography. The column was then equilibrated with several bed volumes of the start buffer, 1.6 M ammonium sulfate. A linear gradient was produced by mixing buffer A (0.1 M phosphate, pH 7.0/2 M ammonium sulfate) and buffer B (0.1 M phosphate alone) starting at 20% buffer B and ending with 80% buffer B. The transferrin eluted at ~40% buffer B, and the albumin eluted at ~45% buffer B. The monoclonal IgG peak eluted at ~55% buffer B and was therefore sharply separated from the contaminating proteins.

Elution by detergents

Addition of detergents to the HIC column is one method for increasing the resolution of elution. However, some detergents can bind very strongly to HIC matrices and are difficult to wash out with common solvents such as ethanol. Detergents should be used with care. The mechanism of detergent action in HIC is that the nonpolar end of the detergent molecule competes with the hydrophobic patches on the protein for the hydrophobic group on the matrix.

An example of detergent elution was published by Hereld et al. (1986), describing the purification of phospholipase C from the protozoan *Trypanosoma brucei*. The enzyme was present in an ammonium sulfate cut prepared from the solubilized membranes of the protozoan. A phenyl column was equilibrated with 25 mM sodium succinate buffer (pH 6.0) containing ammonium sulfate at 50% saturation. The sample (already in ammonium sulfate) was applied to the column and eluted in a descending salt gradient starting with the 50% saturated ammonium sulfate in buffer A and ending with 0% salt (i.e., buffer only). The phospholipase was then eluted by application of buffer containing 1% CHAPS. Most of the protein eluted in a sharp peak immediately after application of the detergent, but there was some trailing in this elution. The yield was 62%, with a 22-fold purification.

Elution by polarity-reducing compounds

The addition of various polarity-reducing compounds to the buffer is another way to increase the resolving power of the elution

PCU IAS/CENTRAL UNIVERSITY LIBRARY

buffer. This can be done only if the protein of interest is stable upon exposure to the compound used. The reagents most commonly used for this purpose are 40% ethylene glycol or 30% isopropanol, dissolved in salt-free buffer. These are used to decrease the polarity of the eluent, resulting in a decrease in binding strength and thus elution of the bound protein from the column.

An example of this technique, using ethylene glycol to elute human pituitary prolactin, was published by Roos et al. (1979). The work was done in the cold, and started with whole organs that were homogenized and extracted at pH 9.8. Between 200 and 400 mg protein was applied to a phenyl column equilibrated with a buffer consisting of 0.2 M glycine adjusted to pH 9.8 with NaOH. Elution was carried out by stepwise reduction of the buffer concentration, first to 0.02 M glycine, pH 9.8, then to 0.002 M glycine, pH 9.8, with 50% (v/v) ethylene glycol added. The last buffer eluted the prolactin. Flow rate for this elution was ~5 cm/hr and the recovery was 95%.

HIC and protein polarity

It has been estimated by Klotz (1970) and Lee and Richards (1971) that 40% to 50% of the surface of a protein is nonpolar. These nonpolar areas are the areas on the protein responsible for binding the protein to hydrophobic matrices in the presence of moderate to high concentrations of salting-out salts. Various hydrophobicity scales for the amino acids have been published; Ericksson (1989) presents several of these in tabular form. There is also a relationship between this salt-promoted interaction with a HIC ligand and the precipitation data for proteins (Melander and Horvath, 1977).

HIC and other chromatographic techniques

HIC is complementary to the other adsorptive techniques, as well as gel filtration, in that it separates on the basis of a very different physicochemical property of the protein than these other methods and can therefore be used in conjunction with them to achieve more complete separation (see *UNIT 8.1*). The physicochemical property exploited by HIC is the presence or absence of the various hydrophobic amino acids in the protein, as distinguished from the property exploited by ion-exchange chromatography, which is the presence or absence of the various charged amino acids. HIC also distinguishes proteins according to the hydrophobic amino acids on the surface of the protein, not those buried inside the folded protein. This distinguishes HIC from reversed-phase chromatography, which also separates on the basis of hydrophobicity, but on the basis of the hydrophobicity of the denatured or unwound sequence. If there is reason to believe it should be possible to exploit the hydrophobicity of the surface of the protein of interest, or to exploit the hydrophobicity of the surface of the contaminants in the sample, then HIC should be included in a separation plan.

Critical Parameters and Troubleshooting

Because the fundamental operating principle of HIC is that hydrophobic interactions are facilitated by high-salt environments, an incomplete equilibration of either the sample or the matrix with the high-salt buffer will have a negative effect on results and reproducibility.

Sample equilibration

The sample must be completely dissolved in the start buffer. This can be done either by buffer exchange through gel-filtration chromatography (*UNIT 8.3*), dialysis (*APPENDIX 3B*), or membrane filtration. Sample solubility should be checked in each salt environment to which the sample will be exposed to during chromatography, to make sure that there will be no precipitation when changing from a high-salt to a no-salt environment.

Column equilibration

The column must be equilibrated in the high-salt environment at the correct pH for sample application. Many chromatographers try short-cuts during column equilibration only to find that they must repeat the run. It is common that at least 10 column volumes of buffer must pass through a column to equilibrate it. The column should be equilibrated at the temperature at which the chromatography will be run. Thus, a column run overnight in the cold room is not equilibrated for a run the next morning at room temperature.

Optimizing resolution

Panels A to F in Figure 8.4.6 show some typical results that may be obtained on different HIC matrices using the methodology of Support Protocol 1. The shaded section of each chromatogram indicates the protein of interest.

Target protein eluted early in the gradient; resolution not satisfactory (Fig. 8.4.6A).

Not much can be gained in this situation by changing the salt concentration of the equili-

PRIVATE/CENTRAL UNIVERSITY LIBRARY

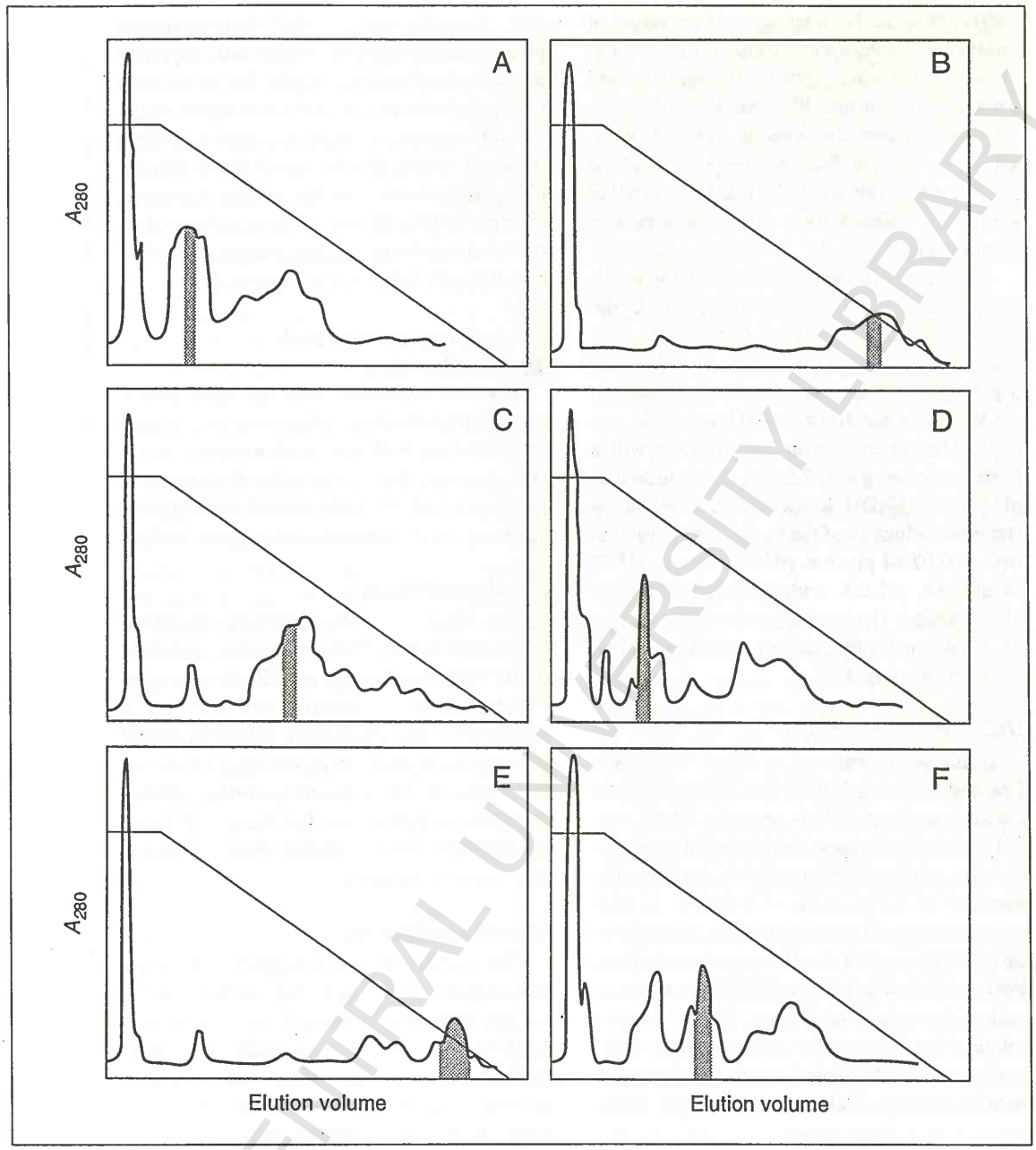


Figure 8.4.6 Chromatograms illustrating various optimal and suboptimal outcomes in HIC. (**A**) Target protein eluted early, resolution not satisfactory. (**B**) Target protein eluted late, resolution not satisfactory. (**C**) Target protein eluted in middle of gradient, resolution not satisfactory. (**D**) Target protein eluted early, resolution satisfactory. (**E**) Target protein eluted late, resolution satisfactory. (**F**) Target protein eluted in middle of gradient, resolution satisfactory (optimal results). These results and techniques for improving upon them are discussed in Critical Parameters and Troubleshooting.

bration buffer. Decreasing the salt concentration will decrease the binding capacity of the protein of interest and might even lead to its elution together with the unbound contaminant fraction. Increasing the salt concentration might lead to adsorption of unwanted impurities and a conseqent decrease in the selectivity of the matrix for the protein of interest.

Changing the pH of the equilibration buffer might result in a stronger binding and higher selectivity for the protein of interest. The effect of pH is variable for different proteins; usually lowering the pH leads to increased binding. Increasing the pH usually leads to a decreased binding of proteins, which in this particular case might result in the elution of the protein of interest together with the unbound fraction.

The appropriate next step in this case would be to try the experiment at a lower and higher

ACU IASTA CENTRAL UNIVERSITY LIBRARY

pH. If no improvement in selectivity is obtained, a matrix with a different ligand or (if available) the same ligand at a higher degree of ligand substitution should be tried.

Target protein eluted near the end of the gradient; resolution not satisfactory (Fig. 8.4.6B).

A decrease in buffer salt concentration will weaken the strength of binding, resulting in earlier elution of the protein of interest. A decrease in salt concentration may also have a positive effect on selectivity because a greater quantity of the less hydrophobic substances will be eluted with the unbound fraction. However, the effect of this approach on the resolution will be marginal, because the contaminants elute very close to the protein of interest—both before and after the target protein peak. Changing the pH of the equilibration buffer may have a positive effect on resolution and should be tried.

The appropriate next step in this separation would be to try the experiment at a lower and higher pH. If no improvement in resolution is obtained, a gel matrix with a different ligand or (if available) the same ligand at a lower degree of ligand substitution should be tried.

Target protein eluted in the middle of the gradient; resolution not satisfactory (Fig. 8.4.6C).

Changing the concentration of the salt in the equilibration buffer will have a limited effect on resolution. However, either an increase or decrease in the pH of the equilibration buffer might have a favorable effect. The appropriate next step in this experiment is to repeat the separation at a higher and lower pH. If no improvement is seen, a gel matrix with a different ligand or (if available) the same ligand at a higher degree of ligand substitution should be tried.

Target protein is eluted early in the gradient; resolution is satisfactory (Fig. 8.4.6D).

The separation shown in panel D is an example of a good choice of matrix for the protein of interest. However, the fact that the protein of interest is eluted very early in the gradient indicates that the binding capacity of the matrix may be low. This may be compensated for, if necessary, by a moderate increase of salt in the equilibration buffer. This may in turn result in a decrease in the selectivity of the adsorbent because some of the unbound proteins might be adsorbed together with the protein of interest. Another negative effect of increased salt concentration may be a decrease in resolution caused by an increase in gradient slope if the total gradient volume or the cycle time is kept constant. Increased salt concentration will also result in increased costs, which may be important in a manufacturing process. Not much can be gained here by changing the pH of the equilibration buffer, as the resolution obtained is satisfactory.

There is no need to try to improve on this separation unless the low binding capacity is a problem. If that is the case and the problems with increased salt concentration mentioned above are encountered, a matrix with a different ligand or (if available) the same ligand at a higher degree of ligand substitution should be tried.

Target protein of interest is eluted near the end of the gradient; resolution is satisfactory (Fig. 8.4.6E).

This also illustrates a good choice of matrix for the protein of interest. Decreasing the concentration of salt in the equilibration buffer will give earlier elution of the protein of interest, reduced cycle time, and decreased cost for salt. A problem in this situation might be that some of the most hydrophobic-contaminating substances are bound so strongly that an organic solvent or chaotropic agent has to be used to clean the column (see Support Protocol 2). Not much can be gained here by changing the pH, because the selectivity is good. If the very strong binding of hydrophobic contaminants poses a problem, a medium with a different ligand or (if available) the same ligand at a lower degree of ligand substitution should be tried.

Target protein elutes in the middle of the gradient; resolution is satisfactory (Fig. 8.4.6F).

The choice of ligand here is good and there is little risk of strong binding of the most hydrophobic contaminants. This panel illustrates an example of good results in hydrophobic-interaction chromatography.

The examples presented in Figure 8.4.6 do not consider two extreme cases that may arise—i.e., that in which the protein of interest is not bound to the column at all, and that in which the protein of interest binds so strongly that it is difficult to elute without the use of denaturing solvents. In both instances a different HIC medium or another chromatographic separation technique should be tried.

PURDUE UNIVERSITY LIBRARY

Although it is assumed for the sake of discussion that the resolution is not satisfactory in some of the chromatograms shown in Figure 8.4.6, the requirements for resolution in any chromatographic step must be stipulated on a case-by-case basis. What seems to be fairly poor resolution at an advanced stage of a separation can often be good enough at an initial capture step where the main objective of the separation is reduction of volume, removal of some critical contaminants, and preparation for further chromatography using a high-resolution matrix.

For more guidelines on troubleshooting chromatographic separations, the Troubleshooting section in *UNIT 8.2* may be consulted.

Anticipated Results

Hydrophobic-interaction chromatography (HIC) is intended for use as one of several chromatographic steps in a protein purification scheme. The product that elutes from a HIC column may not be chemically pure, as HIC will copurify proteins of similar hydrophobicity. Yields will be high (step yields of 60% to 90% are common), and because the conditions of HIC are nondenaturing, activity recovery will be high as well.

Time Considerations

Chromatography using prepacked HIC columns can be run in ~30 min to 4 hr. It will take ~8 hr to pack and equilibrate a column, then load and elute a sample. Cleaning, regeneration, and reequilibration can be carried out overnight. Samples can often be stored overnight in the high-salt start buffers, but the possibility of precipitation should be ruled out first.

Literature Cited

Englard, S. and Seifter, S. 1990. Precipitation techniques. *Methods Enzymol.* 182:285-300.

Ericksson, K.-O. 1989. Hydrophobic interaction chromatography. *In* Protein Purification: Principles, High Resolution Methods and Applications. (J.-C. Janson and L. Rydén, eds.) pp. 207-226. VCH Publishers, New York.

Franks, F. 1988. Characterization of Proteins. *In* Characterization of Proteins (F. Franks, ed.) p. 53. Humana Press, Clifton, N.J.

Hereld, D., Krakow, J.L., Bangs, J.D., Hart, G.W., and Englund, P.T. 1986. A phospholipase C from *Trypanosoma brucei* which selectively cleaves the glycolipid on the variant surface glycoprotein. *J. Biol. Chem.* 261:13813-13819.

Hjerten, S. 1977. Fractionation of proteins by hydrophobic interaction chromatography, with reference to serum proteins. *In* Proceedings of the International Workshop on Technology for Protein Separation and Improvement of Blood Plasma Fractionation. pp. 410-421. Reston, Virginia.

Hodder, A.N., Aquilar, M.I., and Hearn, M.T.W. 1989. High performance liquid chromatography of amino acids, peptides and proteins LXXXIX. The influence of different displacer salts on the retention properties of proteins separated by gradient anion exchange chromatography. *J. Chromatogr.* 476:391-411.

Klotz, I.M. 1970. Comparison of molecular structures of proteins: Helix content, distribution of apolar residues. *Arch. Biochem. Biophys.* 138:704-706.

Lau, K.H., Freeman, T.K., and Baylink, D.J. 1987. Purification and characterization of an acid phosphatase that displays phosphotyrosyl-protein phosphatase activity from bovine cortical bone matrix. *J. Biol. Chem.* 262:1389-1397.

Lee, B. and Richards, F.M. 1971. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* 55:397-400.

Melander, W. and Horvath, C. 1977. Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series. *Arch. Biochem. Biophys.* 183:200-215.

Porath, J., Sundberg, L., Fornstedt, N., and Olsson, I. 1973. Salting-out in amphiphilic gels as a new approach to hydrophobic adsorption. *Nature* 245:465-466.

Porath, J. 1986. Salt promoted adsorption: Recent developments. *J. Chromatogr.* 376:331-341.

Roos, P., Nyberg, F., and Wide, L. 1979. Isolation of human pituitary prolactin. *Biochim. Biophys. Acta.* 588:368-379.

Scopes, R.K. 1994. Protein Purification: Principles and Practice, 3rd ed. Springer-Verlag, New York.

Segel, I.H. 1976. Biochemical Calculations. John Wiley & Sons, New York.

Sparrman, M. 1986. Purification of monoclonal antibodies by hydrophobic interaction chromatography. Presented at the 6th International Symposium on HPLC of Proteins, Peptides, and Polynucleotides, Baden-Baden, Germany.

Srinivasan, R. and Ruckenstein, E. 1980. Role of physical forces in hydrophobic interactions chromatography. *Sep. Purif. Methods* 9:267-370.

Tiselius, A. 1948. Adsorption separation by salting out. *Arkiv för Kemi, Mineralogi, Geologi* 26B:1-5.

Visser, J. and Strating, M. 1975. Separation of lipoamide dehydrogenase isoenzymes by affinity chromatography. *Biochim. Biophys. Acta.* 384:69-80.

von der Haar, F. 1976. Fractionation of proteins by fractional interfacial salting out on unsubstituted agarose gels. *Biochem. Biophys. Res. Commun.* 70:1009-1013.

UCLA IMMUNOGENETICS CENTER LIBRARY

Key Reference

Ericksson, 1989. See above.

Describes both the theory and practice of hydrophobic-interaction chromatography and includes a number of references to the theoretical as well as the applications literature.

Contributed by Robert M. Kennedy
Pharmacia Biotech
Piscataway, New Jersey

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Chromatofocusing

Chromatofocusing (CF) is a versatile, high-resolution chromatographic technique suitable for separating proteins on the basis of their isoelectric points. It may be used to purify proteins with isoelectric points in the range of 3 to 11. Proteins elute from a CF column in descending order of isoelectric point. The unique aspect of this technique is the formation of a descending linear pH gradient in situ by elution with a single buffer containing a large number of buffering species (i.e., the CF elution buffer described here, which contains polybuffer)—the same buffering system used in isoelectric focusing (*UNITS 10.2 & 10.4*). This methodology is in contrast to the mixing of a high- and low-pH buffer prior to the column—the technique used to form pH gradients for elution in ion-exchange chromatography (see *UNIT 8.2*).

The Strategic Planning section discusses choice of chromatofocusing medium and buffers, determination of the quantity of medium needed, and preparation of CF buffers. The Basic Protocol provides the procedure for a typical chromatofocusing experiment, Support Protocol 1 describes regeneration of the chromatofocusing column, and Support Protocol 2 presents several methods for separation of the polybuffer from the purified protein. A schematic representation of the steps in a CF experiment is provided in Figure 8.5.1.

STRATEGIC PLANNING

Selecting CF Medium and CF Buffers

The first step in selecting a medium, start buffer, and elution buffer for chromatofocusing (CF) is to determine the appropriate pH gradient range on the basis of the isoelectric point of the target protein. After the pH gradient range has been selected, Tables 8.5.1 and 8.5.2 may be used to select the appropriate CF gel medium and buffers. Finally, after the CF buffers and CF gel medium have been selected, Table 8.5.1 may be used to determine the amount of buffer required and Table 8.5.2 may be used to determine the amount of gel required for the application.

The CF media and buffers listed in Table 8.5.1 are designed to produce a gradient over a maximum interval of 3 pH units or a minimum of 1 pH unit. For initial experiments a pH gradient spanning 3 pH units should be used. Once the chromatographic behavior of the target protein is characterized, narrower gradients can be used to optimize resolution.

If the isoelectric point of the protein is known, the pH range of the CF experiment is chosen so that the target protein will elute after one-third to one-half of the pH gradient. For example, if the isoelectric point of the target is 6.2, a pH range of 4 to 7 or 5 to 8 can be chosen for initial experiments.

If the isoelectric point of the target protein is unknown, it may be determined by isoelectric focusing (*UNIT 10.2*) or by the test tube method for determining starting conditions for ion-exchange chromatography (*UNIT 8.2*). For an example of the latter method, see Figure 8.2.3, Panel A, where the isoelectric point of the protein is between 5.5 and 7.0. Alternatively, because most proteins have an isoelectric point in the range of 4 to 7 (Gianazza and Righetti, 1980), it is usually appropriate to simply select a pH gradient in that range for initial experiments.

Polybuffer 96 and Polybuffer 74 (Pharmacia Biotech) are amphoteric buffers designed for chromatofocusing in the ranges of pH 6 to pH 9 and pH 4 to pH 7, respectively. They contain a large number of buffering species to provide an even buffering capacity, and

Copyright © 1995 by John Wiley & Sons, Inc.

BCU RSI / CENTRAL UNIVERSITY LIBRARY

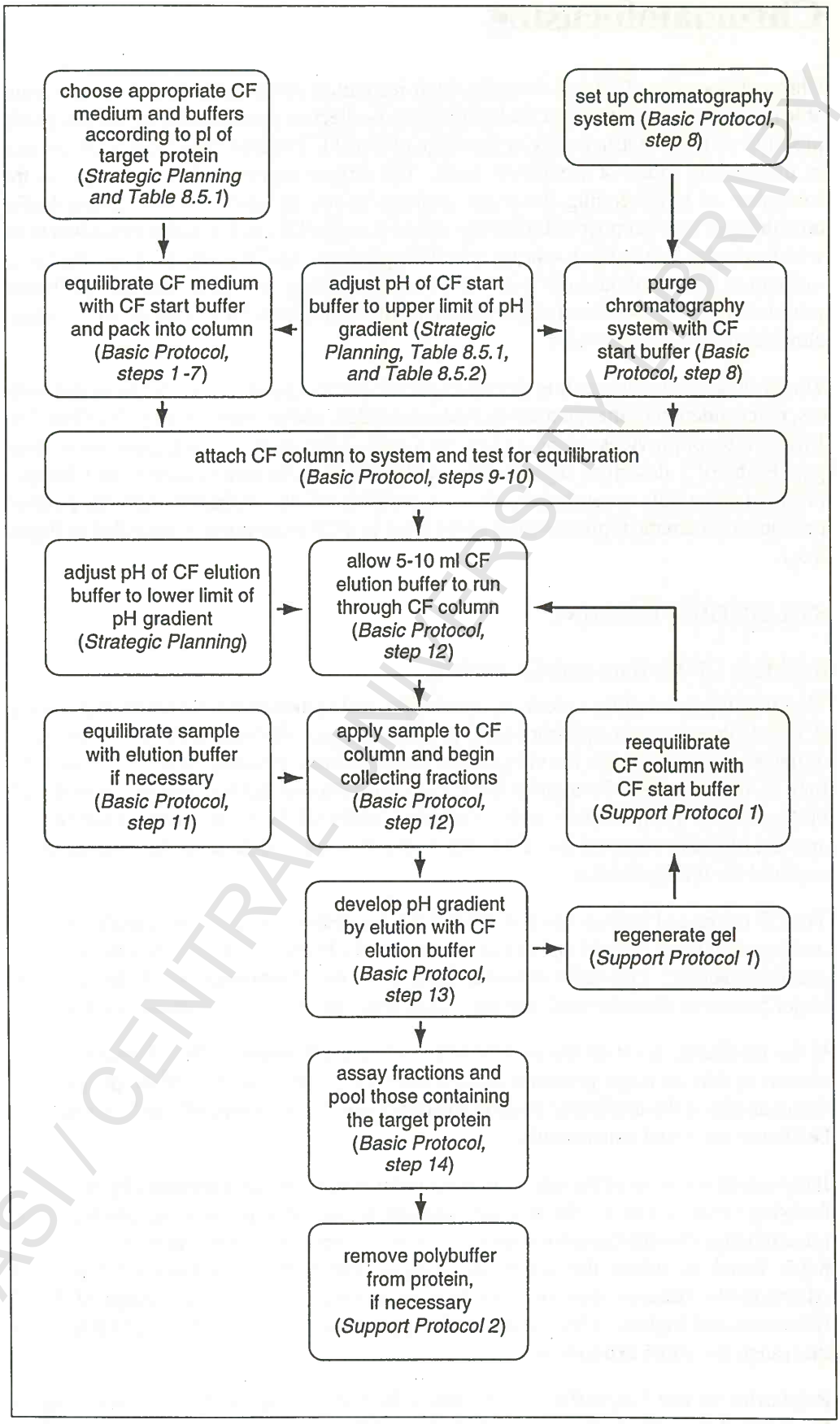

choose appropriate CF medium and buffers according to pI of target protein (Strategic Planning and Table 8.5.1)
set up chromatography system (Basic Protocol, step 8)
equilibrate CF medium with CF start buffer and pack into column (Basic Protocol, steps 1-7)
adjust pH of CF start buffer to upper limit of pH gradient (Strategic Planning, Table 8.5.1, and Table 8.5.2)
purge chromatography system with CF start buffer (Basic Protocol, step 8)
attach CF column to system and test for equilibration (Basic Protocol, steps 9-10)
adjust pH of CF elution buffer to lower limit of pH gradient (Strategic Planning)
allow 5-10 ml CF elution buffer to run through CF column (Basic Protocol, step 12)
equilibrate sample with elution buffer if necessary (Basic Protocol, step 11)
apply sample to CF column and begin collecting fractions (Basic Protocol, step 12)
reequilibrate CF column with CF start buffer (Support Protocol 1)
develop pH gradient by elution with CF elution buffer (Basic Protocol, step 13)
regenerate gel (Support Protocol 1)
assay fractions and pool those containing the target protein (Basic Protocol, step 14)
remove polybuffer from protein, if necessary (Support Protocol 2)

Figure 8.5.1 Flow chart for a chromatofocusing experiment.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

are used together with the polybuffer exchange gel PBE 94 or the prepacked Mono P column (both available from Pharmacia Biotech). For chromatofocusing in the range pH 8 to 11, Pharmalyte pH 8 to 10.5 is used in conjunction with the polybuffer exchange gel PBE 118 (both items also available from Pharmacia Biotech). Suitable CF start buffers for different selected pH gradient ranges are indicated in Table 8.5.1. The normal counter-ion for these anion-exchange media is chloride. Monovalent anions other than chloride may be used as counter-ions, but it is critical that these should have a pK_a at least 2 pH units below the lowest point of the gradient chosen (Hutchens, 1989).

As bicarbonate ions and atmospheric carbon dioxide can cause fluctuations in the pH gradient, all buffers must be degassed before use. Atmospheric carbon dioxide may cause a plateau in the region of pH 5.5 to 6.5, depending on the conditions. These effects are most marked with Polybuffer 96 in pH gradients ending at pH 6, and can be avoided by using acetate as the counter-ion (Pharmacia Biotech, 1985). Alternatively, the lower limit for gradients with Polybuffer 96 may be set at pH 6.5. Acetate is not usually recommended as a counter-ion with Polybuffer 74 because of its higher pK_a. Multivalent anions are not recommended as counter-ions.

Determination of Appropriate Quantity of CF Medium

The amount of gel required for a chromatofocusing run depends on the amount of sample, the nature of the sample and contaminants, and the degree of resolution required. For most separations, a bed volume of 20 to 30 ml is sufficient for separation of samples containing from 1 to 200 mg of protein per pH unit covered by the gradient (e.g., 600 mg of protein for a 3-pH-unit gradient). However, the precise amount of gel used will depend on the resolution required from the experiment, a factor that should be considered during optimization.

Preparation of CF Buffers

CF start buffer is prepared as in Table 8.5.1. CF elution buffers are prepared by diluting Polybuffer 96 or 74 or Pharmalyte 8 to 10.5 with distilled water according to the dilution factors listed in Table 8.5.1. The pH of the CF elution buffer should then be adjusted to the lower limit of the selected pH gradient. Approximately 15 V_c of CF start buffer and 15 V_c of CF elution buffer per run are typically required. This should be enough for equilibration (system, column, and sample), separation, and regeneration.

All buffers should be degassed under vacuum or by sonication prior to use. Buffers to be used with the prepacked column Mono P HR 5/20 should be filtered through a 0.45-µm filter membrane prior to use.

SEPARATION OF PROTEINS BY CHROMATOFOCUSING

BASIC PROTOCOL

In this procedure, the CF column is equilibrated in CF start buffer at a pH higher than that of the CF elution buffer. As the lower-pH CF elution buffer moves through the column, the ion-exchange groups of the medium in the column are titrated by the CF elution buffer, forming a descending linear pH gradient in the column. In a case where the isoelectric point (pI) of a protein is below the pH at which the column is equilibrated, the protein will have a net negative charge at the beginning of the experiment and therefore will bind to the anion-exchange medium in the CF column. As the descending pH gradient slowly forms in the column, the pH at the position where the protein is bound will eventually drop below the pI of the protein, and the net charge on the protein will become positive. The positively charged protein will then desorb from the anion-exchange medium and be carried through the column in the mobile phase. As the desorbed and now positively charged protein advances through the column in the mobile phase, it will pass through

PUBLIC CENTRAL UNIVERSITY LIBRARY

Table 8.5.1 Buffers and Gels for Chromatofocusing in Different pH Ranges[a,b]

Selected pH range	CF gel[c]	Start buffer	Elution buffer	Dilution factor[d]	Approx. volume of buffer needed (column volumes)		
					Dead vol	Gradient vol[e]	Total vol
10.5-9[f]	PBE 118	—	—	—	—	—	—
10.5-8	PBE 118	0.025 M triethyl-amine·HCl, pH 11	Pharmalyte pH 8-10.5 (pH 8.0)[g]	1:45	1.5	11.5	13
10.5-7	PBE 118	0.025 M triethyl-amine·HCl, pH 11	Pharmalyte pH 8-10.5 (pH 7.0)[g]	1:45	2	11.5	13.5
9-8[h]	PBE 94 or Mono P	0.025 M ethanol-amine·HCl, pH 9.4	Pharmalyte pH 8-10.5 (pH 8.0)[g]	1:45	1.5	10.5	12
9-7	PBE 94 or Mono P	0.025 M ethanol-amine·HCl, pH 9.4	Polybuffer 96 (pH 7.0)[g]	1:10	2	12	14
9-6	PBE 94 or Mono P	0.025 M ethanolamine acetate, pH 9.4	Polybuffer 96 (pH 6.0)[i]	1:10	1.5	10.5	12
8-7	PBE 94 or Mono P	0.025 M Tris.Cl, pH 8.3	Polybuffer 96 (pH 7.0)[g]	1:13	1.5	9	10.5
8-6	PBE 94 or Mono P	0.025 M Tris acetate, pH 8.3	Polybuffer 96 (pH 6.0)[i]	1:13	3	9	12
8-5[j]	PBE 94 or Mono P	0.025 M Tris acetate, pH 8.3	30%/70% (v/v) Polybuffer 96/ Polybuffer 74 (pH 5.0)[i]	1:10	2	8.5	10.5
7-6	PBE 94 or Mono P	0.025 M imidazole acetate, pH 7.4	Polybuffer 96 (pH 6.0)[i]	1:13	3	7	10
7-5	PBE 94 or Mono P	0.025 M imidazole·HCl, pH 7.4	Polybuffer 74 (pH 5.0)[g]	1:8	2.5	11.5	14
7-4	PBE 94 or Mono P	0.025 M imidazole·HCl, pH 7.4	Polybuffer 74 (pH 4.0)[g]	1:8	2.5	11.5	14
6-5	PBE 94 or Mono P	0.025 M histidine·HCl, pH 6.2	Polybuffer 74 (pH 5.0)[g]	1:10	2	8	10
6-4	PBE 94 or Mono P	0.025 M histidine·HCl, pH 6.2	Polybuffer 74 (pH 4.0)[g]	1:8	2	7	9
5-4	PBE 94 or Mono P	0.025 M piperazine·HCl, pH 5.5	Polybuffer 74 (pH 4.0)[g]	1:10	3	9	12

[a]Table reproduced with permission of Pharmacia Biotech.

[b]Degas all buffers before use.

[c]PBE 94, Mono P, Pharmalyte 8-10.5, Polybuffer 96, and Polybuffer 74 are products of Pharmacia Biotech (see *SUPPLIERS APPENDIX*).

[d]The dilution factor given is not critical; optimal conditions will be found by experience.

[e]Gradient volume will vary with the exact conditions chosen.

[f]Gradients ending at pH 9 are not recommended because pH 9 is above the pH of Pharmalyte pH 8-10.5.

[g]pH adjusted with HCl.

[h]PBE 118 and Pharmalyte 8-10.5 as well as PBE 94 and Polybuffer 96 (all from Pharmacia Biotech) also cover this range.

[i]pH adjusted with acetic acid.

BOUAS / CENTRAL UNIVERSITY LIBRARY

Table 8.5.2 pH Ranges, Loading Capacities, and Flow Rates for Chromatofocusing Gel Media[a]

CF gel medium	pH range	Recommended loading capacity	Recommended flow rate
PBE 118	8-11	10 mg protein/ml gel	50 cm/hr
PBE 94	4-9	10 mg protein/ml gel	50 cm/hr
Mono P HR 5/20	4-9	25 mg protein/column 5 mg protein/peak	0.5-1 ml/min

[a]Recommended flow rates and loading capacities are for initial experiments; see *UNIT 8.2* for optimization procedures. A peristaltic pump is generally adequate for packing and operation of PBE 94 and PBE 118 columns; the Mono P HR 5/20 column is designed for use at higher pressures and should be used with HPLC or FPLC systems.

the gradient, in which the pH in the column is increasing. When the pH in the column finally exceeds the pI of the protein, the net charge of the protein again becomes negative; hence the protein is again adsorbed to the anion-exchange medium. The net effect, as the protein adsorbs and desorbs many times, is a focusing effect which results in zone sharpening, sample concentration, and very high resolution.

Materials

CF gel medium PBE 94 or PBE 118 *or* prepacked CF column Mono P (Pharmacia Biotech; see Strategic Planning, Table 8.5.1, and Table 8.5.2)
CF start buffer (see Strategic Planning and Table 8.5.1)
CF elution buffer (see Strategic Planning and Table 8.5.1)
Protein sample to be purified

Sintered-glass filter funnel, medium porosity
Side-arm flask
Chromatographic system including:
Appropriate chromatographic column
Fluid delivery system with appropriate pump
Sample injection valve
UV monitor suitable for use at 280 nm
pH monitor with flowthrough electrode (optional)
Recorder
Fraction collector

Additional reagents and equipment for checking packed chromatographic bed (*UNIT 8.4*), desalting via gel-filtration chromatography (*UNIT 8.3*), and protein assay (Chapter 3)

Pack the column

1. Pour the appropriate amount of CF gel medium into a sintered-glass funnel attached to a side-arm flask. Using aspiration, pass start buffer through the gel in the funnel until the pH of the effluent is the same as that of the start buffer.

   Equilibration usually requires 10 to 15 bed volumes of start buffer. The gel should be stirred gently from time to time to ensure complete equilibration.

   Steps 1 to 7 are omitted where a prepacked CF column is used. CF columns may also be equilibrated after packing. If a column is used in which the gel medium was not equilibrated before packing, or if a prepacked column is used, the column should be washed with 2 to 3 V_c of 1 M NaCl to remove any bound substances, then with 10 to 15 V_c of CF start buffer until equilibrated. The column has reached equilibrium when

BCU LIBRARY / CENTRAL UNIVERSITY LIBRARY

the pH and conductivity of the eluant match the corresponding values for the CF start buffer. When the column is equilibrated, proceed to step 8.

2. Disperse gel in a small amount of start buffer (~50% of settled bed volume to be packed) to make a slurry. Degas under vacuum.

3. Mount chromatographic column vertically and flush any air bubbles from underneath the bed support nets. The column outlet should be closed at this point.

   See Figure 8.3.1 for an illustration of the chromatographic column. Best results are obtained on long, narrow columns. For high resolution in preparative applications, a bed height ≥30 cm is desirable. A column diameter of 2.5 cm at a 40-cm bed height (V_c = 200 ml) is usually sufficient for separating ≤2000 mg of protein.

4. Place 2 to 3 ml start buffer in the column and pour gel slurry down a glass rod into the column, taking care to avoid introducing air bubbles.

   If the volume of the slurry is greater than that of the column, a packing extension should be used (see Fig. 8.3.1).

5. Open column outlet and allow gel to settle into column. When all of the suspension has run into the column, remove packing extension (if used) and connect top piece (i.e., flow adaptor; see Fig. 8.3.1) to the column, taking care to exclude air bubbles. Connect column to pump.

6. Run start buffer through column at a flow rate of 100 cm/hr until the gel bed has completely settled.

   Equilibration of a column that has already been packed without previous equilibration can be carried out at a lower flow rate. The column has been equilibrated when the pH and conductivity of the eluant match the corresponding values for the start buffer (usually after 10 to 15 bed volumes of start buffer have passed through the column; see annotation to step 1).

7. Check the quality of the column packing (see *UNIT 8.4*).

   A well-packed column is essential to get the full benefits of chromatofocusing—i.e., the achievement of extremely narrow zones (0.02 pH units). Care must be taken in the choice of dye substance for checking beds, as they may be strongly charged. Bovine cytochrome c (pI = 10.5) is recommended as the dye marker for visual inspection of the bed in CF columns. Dilute acetone solutions are appropriate for efficiency tests (see UNIT 8.4).

Prepare the system

8. Set up the chromatography system according to the manufacturer's instructions without the column in-line, but with the UV monitor in-line. Purge the system of air bubbles by washing with CF start buffer. Check for leakage.

9. Install the packed column in the system and continue to equilibrate with start buffer at the flow rate appropriate for the column (see Table 8.5.2).

   Do not exceed 75% of the maximum flow rate used in packing the column (i.e., 75 cm/hr for 100 cm/hr packing flow rate; see step 6).

10. Check that the UV monitor signal is stable and that the recorder is operational. Collect one fraction and measure the pH, which should be the same as that of the CF start buffer. If the pH is not the same, continue to equilibrate the system by running start buffer through the column.

Prepare sample and apply to column

11. Equilibrate protein sample in CF elution buffer or start buffer by desalting via gel-filtration chromatography (*UNIT 8.3*).

BCU IASI (CENTRAL) UNIVERSITY LIBRARY

Although the best way to ensure that the sample will bind to the column is by buffer exchange into start buffer, this step is only necessary when the ionic strength is high (typically >50 mM). Diluting the sample with distilled water can also be used to decrease the ionic strength. Sample pH is not critical if the buffering capacity of the sample is less than that of the start buffer.

The amount of sample to be applied to the CF column depends on the amount of protein present in each zone. As a guideline for initial experiments, 10 mg protein can be applied for every 1 ml CF medium. This quantity will vary according to the number of proteins present in the sample. The volume of sample is unimportant so long as all of the sample is applied before the substance of interest is eluted from the column. The composition of the sample is important: it should not contain large amounts (>0.05 M) of salt.

12. Using a linear flow rate of 50 to 60 cm/hr, run 5 ml CF elution buffer through the equilibrated CF column, then run the sample in CF elution buffer or CF start buffer.

    Running elution buffer through the column before running the sample insures that sample proteins are never exposed to extremes of pH.

    NOTE: *Do not monitor protein at 254 nm, because CF elution buffers absorb at that wavelength.*

13. Still using a linear flow rate of 50 to 60 cm/hr, begin development of the pH gradient by eluting the sample with CF elution buffer. Begin fraction collection, and continue elution with CF elution buffer until the pH of the fractions collected reaches the desired pH (i.e., the pH of the CF elution buffer).

    When exploring conditions for a new separation, collect 20 fractions over the linear pH gradient. Once the separation has been refined, smaller fractions can be collected to optimize resolution; alternatively the peak collection feature of an automated fraction collector may be employed.

    There is a 1.5 to 3.0 V_c delay before the pH gradient begins to elute from the column. The exact delay volume depends on the buffering capacities of the CF column and CF elution buffer.

14. Assay the fractions collected (see Chapter 3) and pool those containing the target protein.

    The Polybuffer 74 or 96 contained in the CF elution buffer does not interfere with most enzyme assays or amino acid analyses (UNIT 3.3). Although CF elution buffer does not interfere with Coomassie blue–based protein assays, it forms a complex with copper ions and interferes with the Lowry protein assay (UNIT 3.4). Where necessary, CF elution buffer may be removed by any of several methods (see Support Protocol 2).

REGENERATION OF THE CF COLUMN

SUPPORT PROTOCOL 1

Additional Materials (also see Basic Protocol)

1 M NaCl

1. Wash column with 2 to 3 V_c of 1 M NaCl to remove bound material.

   Strongly bound material can often be removed with 0.1 M HCl or 0.1 M NaOH, but the gel should be equilibrated to a pH of 4 to 9 as soon as possible after treatment with these acid or alkaline reagents.

2. Equilibrate column with 10 to 15 bed volumes of CF start buffer in preparation for the next chromatographic run.

   If column is being prepared for storage, equilibrate with 2 or 3 column volumes of 20% ethanol instead of start buffer.

BCU ASI / CENTRAL UNIVERSITY LIBRARY

REMOVAL OF POLYBUFFER FROM PROTEIN

There are several methods for separating polybuffer from proteins where it is necessary to do this in preparation for analysis or further purification, including precipitation, hydrophobic interaction chromatography (HIC), gel filtration, and affinity chromatography.

Precipitation

The simplest method for separating protein from polybuffer is precipitation with ammonium sulfate (Scopes, 1994). Solid ammonium sulfate is added to the relevant fractions until a suitable concentration (80% to 100% saturated) is obtained. The sample is then allowed to stand 1 to 2 hr until the protein precipitates. This procedure is usually relatively simple, as the pH of the fraction is at or near the isoelectric point of the protein. Alternatively, the fractions of interest can be placed in dialysis tubing and dialyzed against saturated ammonium sulfate. Precipitation has the advantage of concentrating and stabilizing the sample, and is also inexpensive.

Hydrophobic Interaction Chromatography (HIC)

HIC (*UNIT 8.4*) can be used in conjunction with ammonium sulfate treatment. Ammonium sulfate is added to the fraction of interest obtained from chromatofocusing to a level of 80% saturation, which will favor hydrophobic interactions. This treatment may result in precipitation of the sample protein. If this happens, the purified sample can be harvested by centrifugation.

Gel Filtration

CF elution buffer can be removed from most proteins by gel filtration (*UNIT 8.3*) on a column packed with Superdex 75 prep grade medium (Pharmacia Biotech). If the molecular weight of the protein is <15,000, some difficulties may be encountered in achieving baseline separation in short columns.

Affinity Chromatography

The protein of interest can be easily be separated from CF elution buffer by taking advantage of the biological activity of the protein. The fractions of interest are passed through a column containing a bioselective adsorbent that binds the desired protein, allowing components of the CF elution buffer to elute unretarded. See Chapter 9 for a variety of affinity chromatography techniques.

COMMENTARY

Background Information

Chromatofocusing was first described by Sluyterman and coworkers (Sluyterman and Elgersma, 1978; Sluyterman and Wijdenes, 1978). They proposed that a pH gradient could be produced on an ion exchanger by taking advantage of the buffering action of the charged group of the ion exchange medium. If a buffer that has been initially adjusted to one pH is run through an ion-exchange column that has been initially adjusted to a second pH, a pH gradient is formed, just as if two buffers at different pH were gradually mixed in the mixing chamber of a gradient maker. If such a pH gradient is used to elute proteins bound to the ion exchanger, the proteins elute in order of their isoelectric points. Furthermore, focusing effects occur, resulting in band sharpening, sample concentration, and very high resolution.

The high resolution attainable with CF makes this technique ideal for the intermediate purification and polishing stages of a purification process (*UNIT 8.1*). The combination of CF with techniques that separate proteins based on physiochemical attributes other than charge characteristics—i.e., HIC (*UNIT 8.4*), gel filtration (*UNIT 8.3*), and affinity chromatography (Chapter 9)—is recommended. If the presence of polybuffer in the sample interferes with the intended use of the target protein, an additional polishing step will be required to remove the polybuffer (see Support Protocol 2).

BCU IAS / CENTRAL UNIVERSITY LIBRARY

Critical Parameters and Troubleshooting

The most critical steps in chromatofocusing are: first, selection of the correct pH gradient range for separation, and then proper matching of the buffering capacity of the CF medium to the buffering capacities of the CF start and elution buffers. Monitoring pH and ionic strength of the column effluent during the run is advised for both the interpretation of results and use in troubleshooting. Conductivity and pH monitors equipped with flow cells may be included downstream from the column; these should be calibrated prior to the experiment. If the conductivity and pH are to be measured in the collected fractions, the measurements should be made as soon as possible after collection to avoid variations from uptake of atmospheric gases.

Linear gradients are required for optimal resolution because nonlinear pH gradients give poor resolution in certain areas of the chromatogram. To produce a linear pH gradient, it is necessary that the CF eluant buffer and CF medium have an even buffering capacity over a wide pH range.

A list of potential problems, causes, and solutions for use in troubleshooting are given in Table 8.5.3.

Anticipated Results

Proteins elute from a CF column in descending order of isoelectric point. The pH at which a protein elutes from the CF column is typically less than its isoelectric point. Recovery from CF in terms of total protein is usually >80%. It is not unusual to notice fluctuations in the pH gradient during the delay in gradient formation (see Basic Protocol, step 14 annotation) or during the elution of concentrated proteins. A

Table 8.5.3 Troubleshooting Guide for Chromatofocusing

Problem	Possible cause	Solution
Protein does not bind to column	Column is incorrectly equilibrated	Check that pH at bottom of column matches that of start buffer
	Start pH is too low	Choose a start buffer with a higher pH
	Ionic strength of eluant is too high	Use buffer composition recommended in Table 8.5.1
	Ionic strength of sample is too high	Remove salt in sample by gel filtration in elution or start buffer (Basic Protocol, step 11)
	Column is contaminated with protein	Regenerate correctly and wash thoroughly before reuse (Support Protocol 1)
Protein binds to but does not elute from column	Eluant composition is incorrect	Check pH of eluant
	pH of eluant is too high	Try a lower pH range
Gradient does not reach desired pH	Eluant composition is incorrect	Check eluant pH; prepare fresh eluant if necessary
	Eluant contains CO_2	Degas all buffers prior to use
Gradient fluctuates	Incorrect counter-ion was used	Use recommended counter-ions (Strategic Planning)
	CO_2 is present	Degas all buffers prior to use; in alkaline regions, degas water before making up eluant
Bands are skewed	Bed net is blocked	Remove bed net; clean or replace with a new net
	Column was poorly packed	Check packing with cytochrome *c* (*UNIT 8.4*); repack bed if necessary.

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IASI / CENTRAL UNIVERSITY LIBRARY

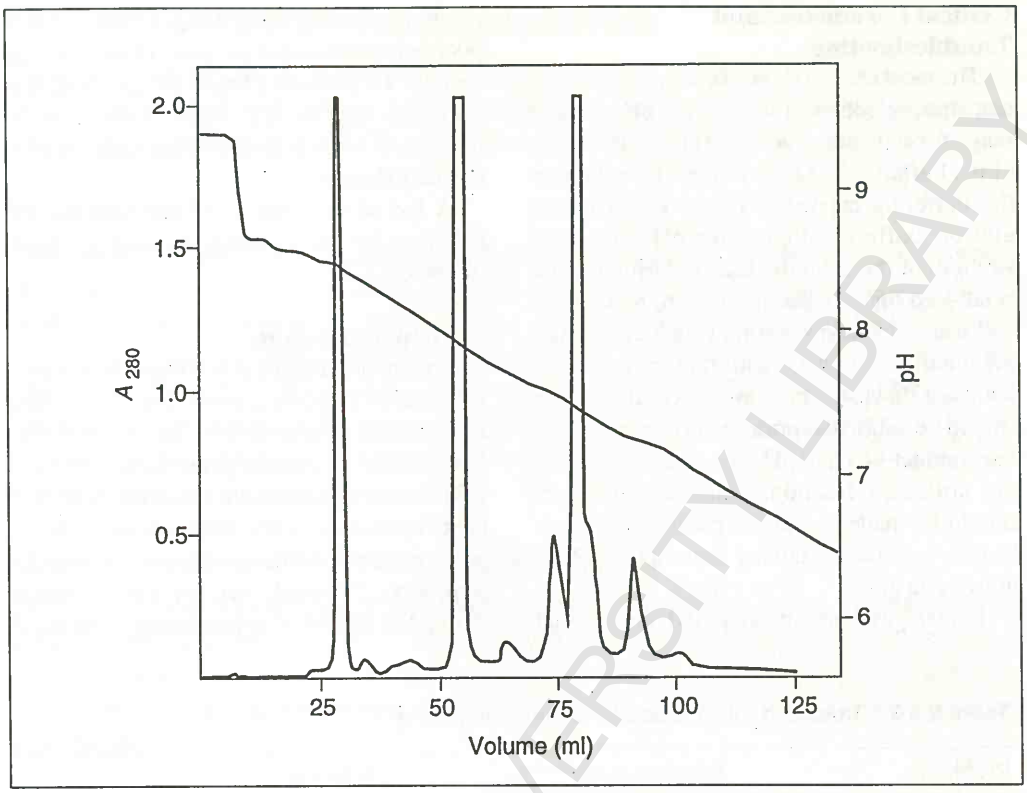


Figure 8.5.2 Chromatogram depicting separation by chromatofocusing with PBE 94 of a model mixture of proteins in the range of pH 9 to 7. An SR 10/50 column (Pharmacia Biotech) with a bed height of 18 cm was used. The sample consisted of 5 ml elution buffer containing 2 mg sperm whale myoglobin, 2 mg horse myoglobin, and 2 mg carboxyhemoglobin. The start buffer consisted of 0.025 M ethanolamine, pH 9.5, and the elution buffer consisted of 0.0075 mmol/pH unit/ml Polybuffer 96, pH 7. A flow rate of 15 cm/hr was used. (Courtesy of Pharmacia Biotech.)

typical chromatogram depicting the separation of a mixture of protein standards by CF is presented in Figure 8.5.2.

Time Considerations

The time required for CF experiments depends primarily on the length of the column, the flow rate, and the volume of sample to be applied; 3 to 4 hr per run is common, excluding the time required for column packing. Separation times using the Mono P HR 15/20 are <1 hr. As with any chromatographic separation, the most time-consuming aspect is normally the analysis of the fractions collected. Increasing resolution with narrow gradients, achieved by optimizing the CF separation, will help reduce the number of fractions to be collected.

Literature Cited

Gianazza, E. and Righetti, P.G. 1980. Size and charge distribution of macromolecules in living systems. *J. Chromatogr.* 193:1-8.

Hutchens, W.T. 1989. Chromatofocusing. *In* Protein Purification: Principles, High Resolution Methods, and Applications (J.C. Janson and L. Rydén, eds.) pp. 149-174. VCH Publishers, New York.

Pharmacia Biotech. 1985. FPLC Ion Exchange and Chromatofocusing: Principles and Methods. Pharmacia Biotech AB, Uppsala, Sweden.

Scopes, R.K. 1994. Protein Purification, Principles and Practice, 3rd ed., pp. 81-85. Springer-Verlag, New York.

Sluyterman, L.A. Æ. and Elgersma, O. 1978. Chromatofocusing: Isoelectric focusing on ion exchange columns. I. General principles. *J. Chromatogr.* 150:17-30.

Sluyterman, L.A. Æ. and Wijdenes, J. 1978. Chromatofocusing: Isoelectric focusing on ion exchange columns. II. Experimental verification. *J. Chromatogr.* 150:31-44.

Key Reference

Pharmacia Biotech. 1984. Chromatofocusing with Polybuffer and PBE. Pharmacia Biotech AB, Uppsala, Sweden.

An excellent discussion of CF theory including practical techniques and model applications for the separation of proteins by CF.

Contributed by Alan Williams
Pharmacia Biotech
Piscataway, New Jersey

BCU IASI CENTRAL UNIVERSITY LIBRARY

Hydroxylapatite Chromatography

Hydroxylapatite (also called hydroxyapatite), a form of calcium phosphate, can be used as a matrix for the chromatography of both proteins and nucleic acids. Molecular separation is not completely dependent on any single property of the molecule such as molecular weight, size, or isoelectric point; the complex interactions between the matrix and the sample confer unique separative properties on hydroxylapatite. It may be used early in purification schedules to fractionate crude mixtures or later on to separate components that have not been well resolved by other steps, such as affinity chromatography or gel filtration.

Two basic protocols are set out in this unit. The first (see Basic Protocol 1) describes standard low-pressure chromatography of a protein mixture using a hydroxylapatite column prepared in the laboratory; the second (see Basic Protocol 2) describes an HPLC method, applicable to proteins and nucleic acids, that uses a commercially available column. Alternate protocols describe column chromatography using a step gradient (see Alternate Protocol 1) or batch binding and step-gradient elution (see Alternate Protocol 2).

STRATEGIC PLANNING

Selecting a Matrix

Choosing the most suitable form of hydroxylapatite matrix for the application is crucial for success. Hydroxylapatite is available in powder form, as a hydrated gel, in a ceramic form, and prepacked into a variety of columns (Table 8.6.1). Most suppliers provide excellent instruction leaflets and/or technical bulletins, and are pleased to offer advice concerning their products. (Also see Critical Parameters.)

Selecting a Column

The optimal size of a hydroxylapatite column is determined by the amount of material to be processed. The binding capacity of the matrix is usually supplied by the manufacturer, and is used to calculate the approximate quantity required. The resin bed should not be too long in relation to its width to allow use of a reasonable flow rate and to minimize broadening of the eluted peaks (see *UNIT 8.1*). A bed height-to-width ratio between 5:1 and 12:1 is suggested. The best type of column to use is one with two adjustable end pieces, so the volume of the bed can be varied as required (see Figure 8.3.1 for illustrations of typical chromatography equipment; also see Critical Parameters).

LOW-PRESSURE CHROMATOGRAPHY FOR PROTEIN PURIFICATION

BASIC PROTOCOL 1

Low-pressure chromatography is frequently used for laboratory-scale purifications (i.e., a few milligrams to a few grams). By choosing a suitably sized column, small or large fractionations can be accommodated.

The procedure described here uses a closed-column pumped system. The matrix is prepared, the column is packed, the sample is chromatographed, and the eluted proteins are visualized by in-line UV detection. Variations on this procedure might include the use of prepacked columns and the detection of eluted proteins by dye binding (*UNIT 3.4*), functional assay, or PAGE (Chapter 10).

Hydroxylapatite can be obtained commercially or, if desired, can be prepared in the laboratory (Atkinson et al., 1973). If the technical literature supplied with the product suggests steps additional to or different from those described in this protocol, always follow the manufacturer's recommendations.

Contributed by Anne V. Broadhurst

Current Protocols in Protein Science (1997) 8.6.1-8.6.12

Copyright © 1997 by John Wiley & Sons, Inc.

BOWMAN-GENTRY UNIVERSITY LIBRARY

Table 8.6.1 Hydroxylapatite Matrices for Standard Chromatography and HPLC

Matrix	Form	Binding capacity (mg/g dry matrix)			Supplier[a]	Remarks/applications
		BSA	DNA	Lysozyme		
Bio-Gel HT	Hydrated	10	>0.5	—	BR	Suitable for general use in purification of proteins and nucleic acids
Bio-Gel HTP	Dry	10	>0.5	—	BR	Dehydrated form of Bio-Gel HT
Bio-Gel HTP (DNA-grade)	Dry	10	>0.8	—	BR	Small particle size to increase capacity and selectivity for dsDNA; suitable for batch use or short columns only
Hydroxylapatite Fast Flow	Dry	>12	>0.7	—	CB	Recommended for applications requiring high flow rate and binding capacity
Hydroxylapatite High Resolution	Dry	>15	>1.2	—	CB	Recommended for analytical applications
Hydroxyapatite Type I	Hydrated	—	—	—	S	
Hydroxyapatite Type III	Hydrated	—	—	—	S	Contains silica gel as crystal initiator
HA-Ultrogel	Hydrated	—	—	—	S	Microcrystalline; requires no precycling; good flow rates even in large columns
Macro-Prep Ceramic Type I	Dry	—	—	>25	BR	Ceramic, macroporous; high protein-binding capacity; best for acidic proteins; excellent flow rates and mechanical stability Macro-Prep
Macro-Prep Ceramic Type II	Dry	—	—	>12.5	BR	Physical properties as Type I; recommended for purification of nucleic acids; low affinity for albamin, therefore recommended for immunoglobulin work
Hydroxylapatite HPLC grade	Dry	>15	>1.2	—	CB	Suitable for HPLC
TSK gel HA-1000	Prepacked column	>20	—	—	TH	Suitable for HPLC; available as prepacked glass and stainless steel columns; use with guard column kit HA-1000 recommended

[a]Abbreviations: BR, Bio-Rad; CB, Calbiochem; S, Sigma; TH, TosoHaas. For addresses and phone numbers of suppliers, see *SUPPLIERS APPENDIX*.

Materials

Hydroxylapatite matrix (see Table 8.6.1)
Loading buffer: 10 mM sodium phosphate, pH 6.8 (*APPENDIX 2E*)
Protein sample to be purified
Gradient solutions: 0.01 and 0.4 M sodium phosphate, pH 6.8 (*APPENDIX 2E*)
Loading buffer (see above) with 0.02% sodium azide

Side-arm flask and vacuum source
Glass chromatography column (see Strategic Planning) with two adaptors or removable end piece and one adaptor (Fig. 8.3.1), tubing, and optional packing reservoir (e.g., Pharmacia Biotech C or XK series)

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Tubing clamps
Peristaltic pump
Buffer reservoirs
Fraction collector
UV monitor with 280-nm filter, connected to a chart recorder
Gradient mixer

Prepare gel

1. Calculate the amount of hydroxylapatite matrix required based on protein-binding capacity.

Capacity of the matrix is usually supplied by the manufacturer

If matrix is supplied dry

2a. Degas a sufficient volume of loading buffer in a side-arm flask under vacuum. Add one part matrix to six parts loading buffer and resuspend by swirling gently.

Do not use a magnetic stirrer as this will damage the hydroxylapatite crystals.

3a. Allow the slurry to settle for a minimum of 20 min, then decant the fines, which are in the cloudy solution above the settled matrix.

4a. Add a volume of loading buffer approximately equal to the volume of the settled bed and mix gently by swirling. Allow matrix to settle.

5a. If the matrix settles to give a sharp dividing line between the bed and clear buffer above, swirl gently to resuspend and proceed to packing of the column (step 6). If the supernatant buffer is still cloudy, repeat steps 3a and 4a, then proceed to step 6.

If matrix is supplied hydrated

2b. Resuspend the matrix (in its own storage buffer as supplied by the manufacturer) by swirling gently.

3b. Pour a sufficient volume of the suspension into a beaker and allow to settle for a minimum of 20 min. Decant the fines.

4b. If the buffer in which the matrix was shipped is different from the loading buffer (10 mM sodium phosphate) add a volume of loading buffer approximately equal to that of the matrix, mix by swirling gently, allow to settle for a minimum of 20 min, then decant.

5b. Add a volume of loading buffer approximately equal to that of the matrix and swirl gently to resuspend. Proceed to packing of the column.

Pack column

6. Examine the column to check that it is clean and undamaged and that sinters, support nets, and O rings are present and intact.

7. Insert bottom adaptor if used, or screw in the end piece. Mount column in a laboratory stand with clamps to ensure that it is vertical. If the volume of the resin slurry is greater than that of the column, fit a packing reservoir (Fig. 8.3.1).

8. Pour a small volume of loading buffer into the column and check that it flows freely through the bottom sinter. Clamp off the outlet tubing with a tubing clamp, leaving ~2 cm of buffer above the sinter.

9. Swirl the hydroxylapatite slurry (step 5a or 5b) gently to resuspend, then pour it all into the column at once and allow 2 to 3 cm to settle under gravity. Release the tubing clamp and allow the rest of the bed to pack under gravity flow.

10. Attach a peristaltic pump to the bottom outlet and pump 5 bed volumes of loading buffer through the column to equilibrate.

PCI LIBRARY / CENTRAL UNIVERSITY LIBRARY

As a rough guide, a standard general-purpose matrix such as Bio-Gel HT (see Table 8.6.1) may be used at a linear flow rate of 75 cm/hr. Fine-particle DNA columns should be run at ~1/10 this rate. Macroporous material—e.g., Macro-Prep Ceramic (Table 8.6.1) can tolerate flow rates up to 5000 cm/hr. The linear flow rate (cm/hr) is equal to the volumetric flow rate (cm^3/hr) divided by the column cross-sectional area (cm^2).

The column should be equilibrated at a flow rate 25% to 50% higher than that anticipated for the separation to ensure that there is no further compaction of the bed during the run.

If the manufacturer recommends a maximum flow rate, this should never be exceeded.

11. Insert the top adaptor, lower it onto the column bed, and tighten.

    Buffer will emerge from the tubing, which should remain filled to prevent air from being drawn into the column.

12. Place the tubing from the top adaptor into a reservoir of loading buffer clamped to the support stand at the same level as the top of the column.

13. Pump a further 5 bed volumes of loading buffer through the column.

    If the bed shrinks so that there is a gap between the top adaptor and the surface of the bed, carefully loosen the top adaptor, lower it onto the bed and retighten. The column is now ready and may be used at once or stored for several weeks at 4°C.

Prepare sample

14. If necessary, free sample of particulate matter by centrifuging or passing it through a 0.45-µm filter.

15. Adjust sample to the same pH and a similar salt concentration as the loading buffer.

    If only the pH of the sample is different from that of the loading buffer, this can be carefully adjusted by adding 10 mM sodium dihydrogen phosphate (to decrease pH) or 10 mM disodium hydrogen phosphate (to increase pH). If the salt concentration is too high, the sample should be dialyzed against an appropriate buffer. Buffer exchange may also be effected by use of a desalting column (UNIT 8.3) or a membrane concentration device (UNIT 4.4). The sample may also be diluted with low-ionic-strength buffer.

Chromatograph sample and regenerate column

16. If the column has been stored for some time, pump a little loading buffer through it and check the bed for air bubbles or cracks and the tubing connections for leaks.

17. Transfer the top adaptor tubing from the buffer reservoir to the sample and pump the sample onto the column, collecting fractions with a fraction collector and monitoring the A_{280} using a UV monitor with a 280-nm filter, connected to a chart recorder.

    Take care not to introduce air into the column.

18. Wash the column with loading buffer until the A_{280} returns to baseline, collecting the flowthrough and wash fraction, as these contain the material which does not bind to the matrix.

19. Degas appropriate volumes of 0.01 and 0.4 M sodium phosphate, pH 6.8, as in step 2a. Connect gradient mixer to column and place the two gradient solutions in the respective reservoirs. Elute the bound material using a linear gradient from 0.01 to 0.4 M sodium phosphate with a gradient volume 6 to 10 times the bed volume. Elute any material still bound by washing with 0.4 M sodium phosphate, pH 6.8. Collect 20 to 40 fractions of a size appropriate to the volume of the gradient until all bound material has been eluted (as judged by the A_{280} of the eluant).

20. Regenerate column by washing with 5 bed volumes of loading buffer, then wash with 2 bed volumes of loading buffer containing 0.02% sodium azide. Store column at 4°C.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CAUTION: *Sodium azide is toxic and potentially explosive. Avoid contact with metals and acids and do not lyophilize solutions containing azide. See* APPENDIX 2A *for additional guidelines on the handling and disposal of sodium azide.*

If the manufacturer recommends a different procedure for column regeneration, this should be followed.

Figure 8.6.1 and Table 8.6.2 show results for low-pressure chromatography.

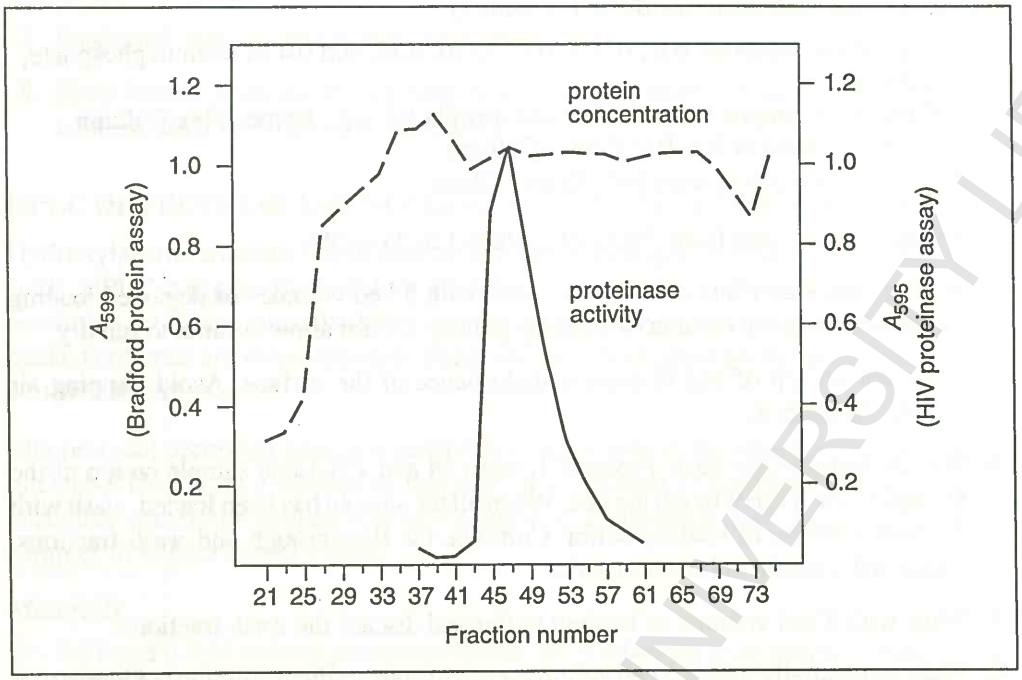


Figure 8.6.1 Chromatogram illustrating purification of HIV-2 proteinase on a hydroxylapatite column (A.V. Broadhurst and A.J. Ritchie, unpub. observ.). 139 g recombinant *E. coli* expressing HIV-2 proteinase were lysed by lysozyme digestion and sonication. The soluble fraction was treated with 1% streptomycin sulfate to precipitate nucleic acids. The supernatant was brought to 25% saturation with solid ammonium sulfate and the pellet was discarded. The ammonium sulfate was increased to 40% saturation and the pellet containing the proteinase was dialyzed against 10 mM sodium phosphate, pH 6.8, containing 0.0125% Tween 20. This fraction was applied to a 4.5 × 19–cm Bio-Rad HT hydroxylapatite column, bed volume 300 ml, at a linear flow rate of 10 cm/hr. The column was washed with 250 ml 10 mM sodium phosphate (pH 6.8)/0.0125% Tween followed by 250 ml 25 mM sodium phosphate (pH 6.8)/0.0125% Tween 20. Bound proteins were eluted with a 600 ml gradient of 25 to 250 mM sodium phosphate, (pH 6.8)/0.0125% Tween. The fractions were assayed for protein by the Bradford method (the dotted plot and the *y* axis on the left; Bradford, 1976; also see *UNIT 3.4*) and for proteinase activity using a colorimetric assay (the solid plot and the *y* axis on the right; Broadhurst et al., 1991).

Table 8.6.2 Typical Purification Scheme for HIV-2 Proteinase Incorporating Hydroxylapatite Chromatography[a]

Purification stage	Total proteinase activity (U)	Specific activity (U/mg)
Crude *E.coli* lysate	79,120	13.9
40% ammonium sulfate pellet	54,590	48.2
Pooled peak fractions from hydroxylapatite column	31,995	194
Pooled concentrated fractions from G-75 gel-filtration column	18,333	561

[a]1 U is arbitrarily defined as the amount of enzyme required to release 1 µmol product/min at 37° in a colorimetric assay (Broadhurst et al., 1991).

BCU LIBRARY CRANFIELD UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 1

SIMPLE COLUMN CHROMATOGRAPHY FOR PROTEIN PURIFICATION USING HYDROXYLAPATITE

If only basic equipment is available, simplified methodology in which the pump is replaced by a gravity column may be applied successfully. The method described below is also useful for quick trial runs to assess the suitability of hydroxylapatite for particular applications.

Additional Materials (also see Basic Protocol 1)

Step gradient solutions: e.g., 0.025, 0.05, 0.10, 0.25, and 0.4 M sodium phosphate, pH 6.8
Column with integral bottom sinter and airtight lid: e.g., Kontes Flex-Column (Kontes Glass) or Bio-Rad Econo-Column
Frit to place on top of resin bed (Kontes Glass)

1. Prepare matrix. (see Basic Protocol 1, steps 1 to 5a or 5b).
2. Pour matrix slurry into column and wash with 5 bed volumes of degassed loading buffer, allowing the column to drain by gravity. Do not allow column to run dry.
3. Place frit on top of bed to prevent disturbance of the surface. Avoid trapping air bubbles beneath it.
4. Prepare sample (see Basic Protocol 1, steps 14 and 15). Load sample on top of the frit and allow it run through the bed. When all the sample has been loaded, wash with 1.5 bed volumes of loading buffer. Combine the flowthrough and wash fractions, which will contain unbound material.
5. Wash with 8 bed volumes of loading buffer and discard the wash fractions.
6. Wash sequentially with 2.5 bed volumes each of step gradient solutions of increasing concentration—e.g., 0.025, 0.05, 0.10, 0.25, and 0.4 M sodium phosphate, pH 6.8 (degassed before use; see Basic Protocol 1, step 2a). Collect each wash separately.

ALTERNATE PROTOCOL 2

BATCH PROCESSING FOR PROTEIN PURIFICATION USING HYDROXYLAPATITE

If hydroxylapatite is to be used with very crude material early in the purification schedule, or if the volume to be processed is very large, this batch separation procedure may be appropriate. Batch processing does not require sophisticated equipment; it can be done quickly and is useful for trial runs.

Materials (also see Basic Protocol 1)

Step gradient solutions: e.g., 0.025, 0.05, 0.10, 0.25, and 0.4 M sodium phosphate, pH 6.8
Sintered-glass funnel
Side-arm flask and vacuum source

NOTE: Degassing of buffers is not essential for batch processing.

1. Prepare matrix (see Basic Protocol 1, steps 1 to 5a or 5b).
2. Carefully remove all supernatant buffer from the settled matrix by decantation.
3. Prepare sample (see Basic Protocol 1, steps 14 and 15).

If the sample volume is small, add sufficient 10 mM sodium phosphate, pH 6.8, to make it approximately twice the bed volume.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

4. Mix the sample with the hydroxylapatite matrix by gentle swirling. Leave for 15 min.
5. Pour slurry into sintered-glass funnel and apply a very gentle vacuum until the buffer just reaches the top of the matrix.

   IMPORTANT NOTE: *Do not allow the matrix to dry out.*
6. Wash matrix in funnel with a small amount of loading buffer. Collect and combine the flowthrough and wash fractions.
7. Wash well with loading buffer and discard wash.
8. Elute bound proteins by washing with a step gradient of increasing phosphate concentration (see Alternate Protocol 1, step 6).

HPLC OF PROTEINS AND NUCLEIC ACIDS USING HYDROXYLAPATITE

BASIC PROTOCOL 2

Hydroxylapatite columns can be used for the chromatography of both proteins and nucleic acids. HPLC is especially useful for analytical applications, as the technique offers high resolution, excellent reproducibility, and fast analysis times. For occasional users, ready-packed columns are recommended. These are usually of good quality and are tested and certified before sale.

The protocol described here is a general one applicable to the separation of proteins or nucleic acids. Using a method of this type, it is possible to quantitate the amount of a particular protein or nucleic acid in a sample by comparison with a range of standard samples of known concentration.

Materials

0.01 and 0.4 M sodium phosphate buffer, pH 6.8 (*APPENDIX 2E*; prepare using boiled H_2O)
Protein or nucleic acid standard solutions
Protein or nucleic acid test samples
0.2 M sodium hydroxide
0.01 M sodium phosphate, pH 6.8 (*APPENDIX 2E*) with 0.02% (w/v) sodium azide
0.45-µm filters (for buffers)
0.22-µm filters (for samples)
~7.5-mm × 7.5-cm HPLC column packed with suitable grade of hydroxylapatite (Table 8.6.1)
Guard column recommended for HPLC column used
HPLC apparatus equipped with pumps, gradient mixer, autosampler or injection loop, integrator, and UV monitor

Prepare buffers and sample

1. Filter 0.01 and 0.4 M sodium phosphate buffers through a 0.45-µm filter. Degas in side-arm flasks under vacuum before use.
2. Filter sample through a 0.22-µm filter.

   If sample is very turbid it may be necessary to centrifuge it first to avoid blocking the filter.

Prepare column and chromatograph samples

3. Follow any instructions supplied by the manufacturer relating to preparation of column and assembly of HPLC system.
4. Equilibrate column with 0.01 M sodium phosphate buffer, pH 6.8, prepared as in step 1.

BOJASI CENTRAL UNIVERSITY LIBRARY

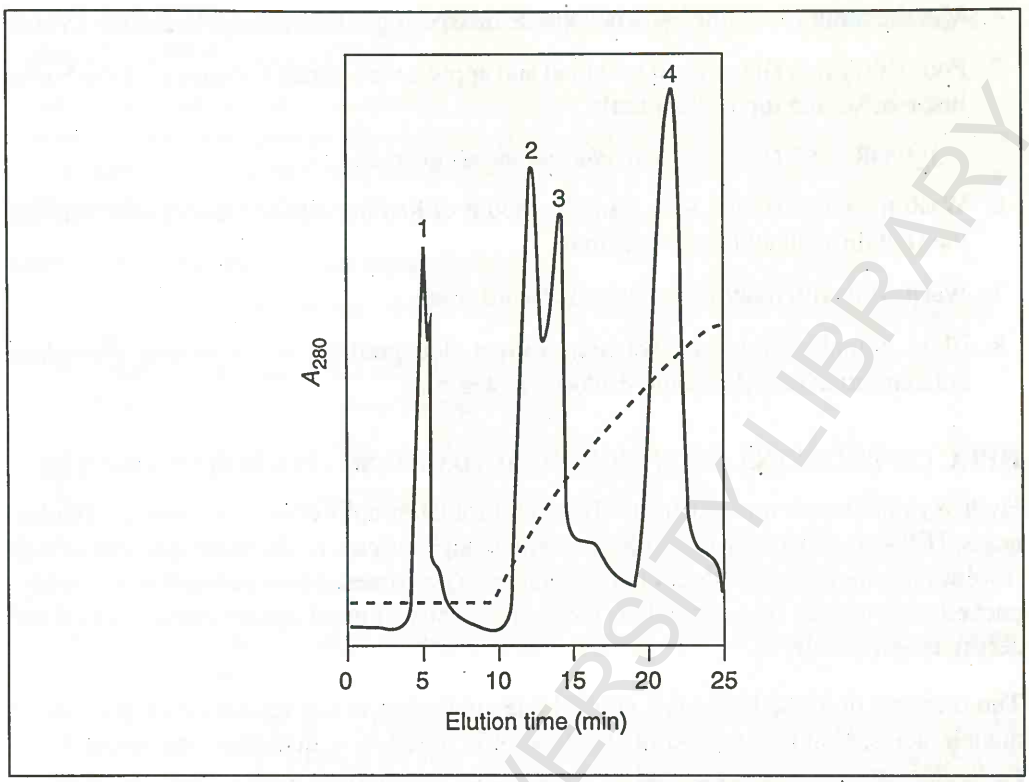


Figure 8.6.2 Chromatogram illustrating purification of monoclonal antibody from mouse ascites fluid. A Bio-Rad Bio-Gel HPHT column with guard column was used and 0.5 ml mouse ascites fluid was applied. The column was eluted with a linear gradient of 0.01 to 0.30 M sodium phosphate, pH 6.8 at a flow rate of 1.0 ml/min. Peak 1 contains low-molecular-weight material, peak 2 contains albumin and serum protease, peak 3 contains other serum (ascites) proteins, and peak 4 contains essentially pure IgG. The linear gradient is traced via the conductance (dotted line). Adapted from Bio-Rad Technical Bulletin US/EG 1115.

5. Program the HPLC to inject 20 µl of sample and set the UV monitor to 260 nm for nucleic acids or 280 nm for proteins. Place 0.01 and 0.4 M sodium phosphate buffers, pH 6.8, in the appropriate reservoirs of the gradient mixer and run a 30-min linear gradient from 0.01 to 0.4 M sodium phosphate at a flow rate of 1 ml/min.
6. Run a standard through the column to to check its performance using the parameters in step 5. Collect 1-ml fractions.
7. Run the standard at a range of concentrations to generate a standard curve.
8. Run the test samples.
9. Use the integrated peak areas of the standards and test samples to calculate the concentration of the protein/nucleic acid of interest in the samples.

See Figure 8.6.2 for representative results.

Regenerate and store column

10. At the end of each sample batch, inject 1 column volume of 0.2 M sodium hydroxide to remove any strongly bound material. Wash extensively with 0.4 M sodium phosphate.
11. Store column at room temperature in 0.01 M sodium phosphate, pH 6.8 containing 0.02% sodium azide.

CAUTION: *Sodium azide is toxic and potentially explosive. Avoid contact with metals and acids and do not lyophilize solutions containing azide. See APPENDIX 2A for additional guidelines on the handling and disposal of sodium azide.*

BCU IASI CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

Calcium phosphate gels have been used for chromatography for many years. The material originally used was unsatisfactory with respect to flow rate, stability, and elution characteristics, and fell out of favor as newer matrices such as HA Ultrogel and Macro-Prep Ceramic became available. In the 1950s, Tiselius and co-workers described a process for preparing hydroxylapatite, a crystalline form of calcium phosphate, and demonstrated its usefulness for protein separation (Tiselius et al., 1956). Material prepared by this method demonstrated superior flow rates and has been commercially available for many years. Recently, some new forms of hydroxylapatite with spherical particles have been developed, and these have demonstrated even better flow rates as well as enough mechanical strength to be used in high-pressure systems. These technological advances, combined with hydroxylapatite's unique mode of separation, make this medium a useful addition to the chromatographer's repertoire.

Protein binding to hydroxylapatite is mediated through interactions between the amino and carboxyl groups of the protein and the Ca^{2+} and PO_4^{3-} groups of the hydroxylapatite crystal lattice. Basic proteins are thought to bind via electrostatic interactions between their amino groups and the surface PO_4^{3-} groups, the positive charge rather than the chemical nature of the group being the important factor (Gorbunoff, 1984b). Acidic and neutral proteins bind to the Ca^{2+} sites via their carboxyl groups; this interaction encompasses both electrostatic interactions and specific effects (Gorbunoff, 1984b). As a general principle, basic proteins with an isoelectric point (pI) of 8 can be eluted by increasing the concentration of the phosphate buffer, and they are also displaced by very low concentrations of Mg^{2+} or Ca^{2+} ions and moderate concentrations of NaCl (Gorbunoff, 1984a). Acidic or neutral proteins can be eluted by increasing the phosphate concentration, but not by adding Mg^{2+} or Ca^{2+} ions or NaCl. Increasing the pH of the elution buffer promotes desorption of proteins irrespective of pI (Gorbunoff, 1984a).

There are many proteins whose interactions with hydroxylapatite are not as easily predicted on the basis of pI alone—e.g., β-lactoglobulin, ovomucoid, trypsinogen (Gorbunoff and Timasheff, 1984), phosphoproteins (Bernardi et al., 1972), and metalloproteins (Atkinson et al., 1973). Tertiary structure is also important for binding, and proteins with floppy or unstructured regions such as the glycoprotein ovomucoid, or those that have been denatured, show reduced affinity for hydroxylapatite (Gorbunoff, 1984a). The distribution of surface charge may also play a part (Gorbunoff, 1984b).

Nucleic acids bind to hydroxylapatite by interaction of the charged phosphate backbone with the Ca^{2+} ions at the crystal face. Double-stranded DNAs bind more tightly than single-stranded DNAs, possibly because their planar nature maximizes the interaction. (Martinson, 1973). Small double- and single-stranded DNAs appear to separate by size. tRNAs can also be separated, and secondary and tertiary structures may be important in this context (Kawasaki et al., 1986).

Hydroxylapatite chromatography is performed using mild conditions at moderate pH and salt concentration, which favors the preservation of biological activity. The matrix has low nonspecific binding, and small molecules such as amino acids do not bind well. It is chemically stable at pHs between 6 and 11 and can be used with denaturing agents, detergents, and reducing agents such as dithiothreitol. EDTA and chelating buffers such as PIPES should be avoided as these bind the calcium.

The unique mode of separation of hydroxylapatite, combined with the improved handling characteristics of the newer forms available, make it a useful adjunct to the more commonly used techniques such as ion-exchange and hydrophobic-interaction chromatography. Hydroxylapatite has proved useful in the separation of IgG idiotypes that cannot easily be separated by other techniques (Juarez-Salinas et al., 1986), as well as for the isolation of immunoglobulins from ascites fluid and tissue culture media (Smith et al., 1984). Another application may be the secondary purification step for recombinant proteins with N-terminal affinity tags such as maltose-binding protein (MBP), glutathione-S-transferase (GST), or 6His. When these proteins have been isolated by affinity chromatography, any breakdown products produced by C-terminal degradation and any premature termination fragments will also bind to the resin, as they are tagged. These contaminants can be difficult to remove by gel filtration or ion-exchange chromatography as they may differ very little in size or charge from the protein of interest; hence hydroxylapatite chromatography may assist in the separation.

B.C.U. „M. EMINESCU" IAȘI
CHI IE

As an example, during the preparation of human cytomegalovirus proteinase for crystallographic studies, protein that had been purified using Mimetic Green and Q Sepharose and that appeared to be homogeneous on PAGE, separated into two peaks on a hydroxylapatite column (A.V. Broadhurst, unpub. observ.).

The ability of hydroxylapatite to bind and separate nucleic acids can be advantageous during chromatography of proteins derived from recombinant *E. coli.* Crude extracts contain large quantities of DNA, which can bind to proteins or ion-exchange resins and contaminate the product. Hydroxylapatite can be used to remove nucleic acids during the purification.

Critical Parameters

Selecting format

Careful consideration must be given to choosing the most appropriate chromatographic format for the application. If hydroxylapatite is to be used with very crude material early in the purification schedule, or if the volume to be processed is very large, a batch separation (see Alternate Protocol 2) may be suitable. Batch processing does not require sophisticated equipment; it can be done quickly and is useful for trial runs.

Low-pressure chromatography (see Basic Protocol 1 and Alternate Protocol 1) can be used to prepare milligram to gram quantities of protein. Semiautomated or fully automated systems are available, but satisfactory results can be obtained with quite basic equipment. This type of chromatography is used primarily for preparative rather than analytical purposes.

HPLC (see Basic Protocol 2), although it can be scaled up for preparative use, is particularly suitable for analytical applications. Its main advantages are high resolution, reproducibility, and rapid throughput.

Selecting matrix

Various grades and types of hydroxylapatite are commercially available (see Table 8.6.1). In general, hydroxylapatite with a small particle size will be suitable for analytical applications, giving excellent resolution but operating at low flow rates. Material with a larger particle size can be used at higher flow rates but with some loss of resolution. Hydroxylapatite can be difficult to work with as its crystalline structure is fragile and easily damaged, but some of the newer types (TSKgel HA-1000 from TosoHaas and MacroPrep ceramic hydroxylapatite from Bio-Rad) are purported to be more robust and are suitable for use with-high pressure systems.

There are several types of ready-prepared columns available commercially. These are designed for particular applications such as HPLC or low-pressure chromatography; some are of particular utility for the separation of proteins and others for separation of nucleic acids.

Selecting buffers

Sodium or potassium phosphate buffers at near-neutral pH are used for the majority of hydroxylapatite applications. Low phosphate concentration and low pH increase protein binding; high phosphate concentration and high pH promote desorption. In this context, low and high pH refer to the range 6 to 8, over which phosphate is an effective buffer. Potassium phosphate is more soluble than sodium phosphate at lower temperatures and is recommended for cold-room use.

The protein-binding capacity of hydroxylapatite is maximal at pH 6.6 to 7 and decreases sharply above pH 7 (Atkinson et al., 1973). The matrix itself is chemically stable over a pH range of 6 to 11, but dissolves below pH 5.5. Buffers other than phosphate have been employed—e.g., MOPS with a phosphate gradient (de Gunzberg et al., 1984)—but should be used with caution as they may damage the matrix. 10 mM Tris·Cl can be used, but may cause some deterioration over time. Chelating buffers such as PIPES must be avoided. Buffer pH should always be within the range over which the sample protein is known to be stable.

Various buffer additives can be employed. Glycerol, added to stabilize proteins, can be used at a concentration of 10% to 15%; it increases buffer viscosity so that a reduction in flow rate may be necessary. Reducing agents such as dithiothreitol and 2-mercaptoethanol, as well as nonionic detergents such as Triton X-100 and Tween 20, are compatible with hydroxylapatite. SDS has been used, but can reduce nonspecific binding and alter elution characteristics; it is also rather insoluble at 4°C and therefore is probably best avoided. Chelating agents such as EDTA are contraindicated as they bind to the calcium ions of hydroxylapatite and will rapidly destroy the column.

Troubleshooting

Many of the problems encountered during hydroxylapatite chromatography are common to all types of column chromatography, but a few are unique owing to the crystalline mineral nature of the matrix.

Column bed drying out or becoming filled with bubbles

Drying out of the bed is usually caused by the system not being airtight. Ensure that all O rings and seals are present and in good condition. It may help to have the buffer reservoir slightly higher than the top of the column. Drying out can also occur if the top sinter or the top of the bed becomes clogged, causing the pump to create a vacuum in the bed. This might be alleviated by reducing the flow rate; if this does not work, remove the top sinter, wash it, and replace. If the problem persists it may be necessary to modify the sample preparation.

Bubbles that appear in the gel are usually caused by inadequate degassing of the buffers. Dissolved gases will be released under reduced pressure or if the buffers/column are moved from a colder to a warmer environment. Degas all buffers (unless they contain volatile solvents) under vacuum before use. Bubbles may also appear if high concentrations of detergent are included in the buffers; this can be minimized by avoiding shaking or vigorous stirring of the buffers and using a moderate flow rate.

Inadequate or decreasing flow rate

This problem may be caused by packing the bed at too high a flow rate, causing it to become compressed. Manufacturers' recommended flow rates should never be exceeded; this is particularly important with hydroxylapatite because of its inherently fragile nature.

The presence of fines (resulting from inadequate removal before packing the bed) or degraded material can also cause poor flow rates. Ensure that the matrix is prepared as recommended; hydroxylapatite is easily damaged by rough treatment such as stirring with magnetic stirring bars and should be resuspended only by gentle swirling.

If hydroxylapatite columns are stored for long periods the top of the bed may become hard and crusted, causing reduced flow rate. This results from the reaction of the calcium ions with carbonate from water used to prepare the buffers. Boiled, degassed water should be used. The hard, crusted material can be removed and replaced with a little fresh hydroxylapatite slurry.

If the flow rate achievable is not high enough to allow the experiment to be performed in a reasonable time, consider whether a matrix with a larger particle size would be more suitable, or whether a batch process might be preferable.

If the flow rate is constant during the preliminary washing and equilibration of the column but decreases during sample loading, the problem may lie with the sample itself. It is important that particulate matter be removed from the sample before loading or it will block the sinter or the top of the gel bed. The sinter can be removed, washed, and replaced as an interim measure, but improved sample preparation is the cure. Crude extracts of recombinant *E. coli*, which are increasingly used as the starting point for protein purification, can be problematic. Even if the extract is clarified by centrifugation, it may be extremely viscous due to the high DNA content. It is often possible to reduce the viscosity enough to achieve a reasonable flow rate by diluting one part sample with three parts buffer before loading. If this fails, try precipitating the nucleic acids with 0.01% aqueous polyethyleneimine or 1% streptomycin sulfate, or treat the sample with DNase. *UNIT 4.5* contains additional guidelines on clarification of samples and nucleic acid removal.

Protein not bound to column

If the protein of interest is only partially bound, with significant amounts appearing in the flowthrough, the binding capacity of the column may have been exceeded. Try increasing the ratio of matrix to total protein.

If the protein is not bound using 10 mM sodium phosphate, pH 6.8, try reducing the molarity of the sodium phosphate. Concentrations as low as 1 mM can be used, although the buffering capacity will be small at this concentration. If the protein is stable at lower pH, the pH of the loading buffer can be decreased to 6 to facilitate binding. Do not reduce the pH below 5.8, as the hydroxylapatite matrix will dissolve at lower pH.

Adding 0.005 mM $CaCl_2$ to the buffer will increase binding of acidic proteins. $CaCl_2$ should only be used with low molarities of phosphate (<200 mM) or calcium phosphate will precipitate out of solution.

Even if none of the above modifications are successful, hydroxylapatite chromatography might still be a useful initial step if a significant amount of contaminating material is removed by binding to the matrix.

Poor separation of proteins on gradient and problems with eluting bound protein

If there are protein peaks that do not separate completely, try increasing the volume of the gradient. Do not exceed more than ~12 bed

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

volumes as there will be no benefit from this.

If all the bound material elutes very early on in the gradient, try using a narrower gradient—e.g.,10 mM to 150 mM sodium phosphate, pH 6.8.

If most of the proteins elute near the end of the gradient, try using a higher loading concentration of sodium phosphate—e.g., 50 to 100 mM.

The molarity of the elution buffer can be increased. If phosphate concentrations >0.4 M are required for complete elution and the chromatography is to be done at 4C, use potassium phosphate buffer instead of sodium phosphate, as it is more soluble. Desorption of proteins can also be achieved by increasing the pH of the buffer.

Basic proteins can be eluted with Ca^{2+} at concentrations of ~0.003 M. This cannot be done in phosphate buffer as calcium phosphate will precipitate out. Wash the column with 1 mM NaCl, then with 0.003 M $CaCl_2$ in 1 mM NaCl.

Anticipated Results

High recoveries can be expected with hydroxylapatite; near-quantitative yields have been reported with 90% to 95% being common. Retention of activity is usually excellent because of the mild conditions employed (Kato et al., 1987). Depending on the nature of the sample to be chromatographed, >20-fold purification is achievable.

Time Considerations

Small columns can be prepared and a chromatographic run carried out in a working day. Larger columns take longer to run and therefore will need to be prepared in advance.

Literature Cited

Atkinson, A., Bradford, P.A., and Selmes, I.P. 1973. The large scale preparation of chromatographic grade hydroxylapatite and its application in protein separation procedures. *J. Appl. Chem. Biotechnol.* 23:517-529.

Bernardi, G., Giro, M.-G., and Gaillard, C. 1972. Chromatography of polypeptides and proteins on hydroxyapatite columns: Some new developments. *Biochim. Biophys. Acta* 278:409-420.

Bio-Rad Technical Bulletin 1115 US/EG. Bio-Gel HPHT for protein and nucleic acid HPLC: New high performance hydroxyapatite column. Bio-Rad Laboratories, Hercules, Calif.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.

Broadhurst, A.V., Roberts, N.A., Ritchie, A.J., Handa, B.K., and Kay, C. 1991. Assay of HIV-1 proteinase: A colorimetric method using small peptide substrates. *Anal. Biochem.* 193:280-286.

de Gunzberg, J., Part, D., Guiso, N., and Veron, G. 1984. An unusual adenosine 3′, 5′-phosphate dependent protein kinase from *Dictyostelium discoideum*. *Biochemistry* 23:3805-3812.

Gorbunoff, M.J. 1984a. The interaction of proteins with hydroxyapatite. 1. Role of protein charge and structure. *Anal. Biochem.* 136:425-432.

Gorbunoff, M.J. 1984b. The interaction of proteins with hydroxyapatite. 2. Role of acidic and basic groups. *Anal. Biochem.* 136:433-439.

Gorbunoff, M.J. and Timasheff, S.N. 1984. The interaction of proteins with hydroxyapatite. 3. Mechanism. *Anal. Biochem.* 136:440-445.

Juarez-Salinas, H., Ott, G.S., Chen, J.-C., Brooks, T.L., and Stanker, L.H. 1986. Separation of IgG idiotypes by high performance hydroxylapatite chromatography. *Methods Enzymol.* 131:615-622.

Kato, Y., Nakamura, K., and Hashimoto, T. 1987. High performance hydroxyapatite chromatography of proteins. *J. Chromatogr.* 398:340-346.

Kawasaki, T., Ikeda, K., Takahashi, S., and Kuboki, Y. 1986. Further study of hydroxylapatite high-performance liquid chromatography using both proteins and nucleic acids, and a new technique to improve chromatographic efficiency. *Eur. J. Biochem.* 155:249-257.

Martinson, H.G. 1973. The basis of fractionation of single-stranded nucleic acids on hydroxylapatite. *Biochemistry* 12:2731-2736.

Smith, G.J., McFarland, R.D., Reisner, H.M., and Hudson, G.S. 1984. Lymphoblastoid cell–produced immunoglobulins: Preparative purification from culture medium by hydroxylapatite chromatography. *Anal. Biochem.* 141:432-436.

Tiselius, A., Hjerten, S., and Levin, O. 1956. Protein chromatography on calcium phosphate columns. *Arch. Biochem. Biophys.* 65:132-155.

Key References

Gorbunoff, 1984a. See above.

Gorbunoff, 1984b. See above.

Gorbunoff and Timasheff, 1984. See above.

A series of three consecutive papers describing the basis of interaction of proteins with hydroxylapatite. Practical aspects are described, and there is extensive discussion of theoretical considerations.

Anne V. Broadhurst
Roche Products Ltd.
Welwyn Garden City, Hertfordshire
United Kingdom

BCU IAS CENTRAL UNIVERSITY LIBRARY

AFFINITY PURIFICATION

9

INTRODUCTION 9.0.1

9.1 Lectin Affinity Chromatography 9.1.1
- Basic Protocol: Con A–Sepharose Affinity Chromatography 9.1.4
- Support Protocol: Pilot Study to Determine Lectin Binding and Elution Conditions 9.1.6
- Alternate Protocol: Wheat Germ Agglutinin (WGA)–Agarose Affinity Chromatography 9.1.7
- Reagents and Solutions 9.1.8
- Commentary 9.1.8

9.2 Dye Affinity Chromatography 9.2.1
- Basic Protocol 1: Selection of Components Using Chromatography 9.2.1
- Alternate Protocol 1: Selection of Components Using Centrifugation 9.2.4
- Basic Protocol 2: Negative Chromatography 9.2.6
- Basic Protocol 3: Positive Chromatography Using Batch Elution 9.2.7
- Alternate Protocol 2: Positive Chromatography Using Gradient Elution 9.2.8
- Basic Protocol 4: Tandem Chromatography 9.2.9
- Support Protocol: Immobilization of Reactive Dyes 9.2.10
- Commentary 9.2.13

9.3 Affinity Purification of Natural Ligands 9.3.1
- Basic Protocol 1: Cyanogen Bromide Activation 9.3.2
- Alternate Protocol 1: Activation with *p*-Nitrophenyl Chloroformate 9.3.5
- Alternate Protocol 2: Activation with Tresyl Chloride 9.3.6
- Reagents and Solutions 9.3.7
- Commentary 9.3.8

9.4 Metal-Chelate Affinity Chromatography 9.4.1
- Strategic Planning 9.4.1
- Basic Protocol: Native MCAC for Purification of Soluble Histidine-Tail Fusion Proteins 9.4.2
- Alternate Protocol 1: Denaturing MCAC for Purification of Insoluble Histidine-Tail Fusion Proteins 9.4.6
- Alternate Protocol 2: Solid-Phase Renaturation of MCAC-Purified Proteins 9.4.8
- Support Protocol 1: Analysis and Processing of Purified Proteins 9.4.9
- Support Protocol 2: NTA Resin Regeneration 9.4.10
- Reagents and Solutions 9.4.11
- Commentary 9.4.12

9.5 Immunoaffinity Chromatography 9.5.1
- Basic Protocol: Isolation of Soluble or Membrane-Bound Antigens 9.5.1
- Alternate Protocol 1: Low-pH Elution of Antigens 9.5.4
- Alternate Protocol 2: Batch Purification of Antigens 9.5.5
- Alternate Protocol 3: Elution in Octyl β-D-Glucoside 9.5.5
- Support Protocol: Preparation of Antibody-Sepharose 9.5.6
- Reagents and Solutions 9.5.8
- Commentary 9.5.9

continued

BCU IAS CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

9.6 Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography **9.6.1**
Basic Protocol 1: Preparation of DNA Affinity Resin 9.6.1
Alternate Protocol: Coupling the DNA to Commercially Available CNBr-Activated Sepharose 9.6.6
Support Protocol 1: Purification of Oligonucleotides by Preparative Gel Electrophoresis 9.6.7
Basic Protocol 2: DNA Affinity Chromatography 9.6.8
Support Protocol 2: Selection and Preparation of Nonspecific Competitor DNA 9.6.10
Reagents and Solutions 9.6.11
Commentary 9.6.13

9.7 Purification of DNA-Binding Proteins Using Biotin/Streptavidin Affinity Systems **9.7.1**
Basic Protocol: Purification Using Batch Method 9.7.2
Alternate Protocol 1: Purification Using a Microcolumn 9.7.4
Alternate Protocol 2: Purification Using Streptavidin-Agarose 9.7.5
Support Protocol: Mobility-Shift Assay 9.7.6
Reagents and Solutions 9.7.10
Commentary 9.7.10

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 9
Affinity Purification

INTRODUCTION

Affinity chromatography in the context of this chapter refers to a variety of techniques that exploit the fact that proteins can act as ligands for a variety of ligand-binding proteins (e.g., antibodies, enzymes, and lectins) or are themselves ligand-binding proteins (i.e., act as receptors). Because receptor-ligand interactions are very specific and receptor and/or ligand sites are often unique to the protein of interest, exploitation of this methodology often yields pure or nearly pure proteins with a minimal number of purification steps.

Some receptor or ligand sites, such as antigenic epitopes and the active sites of enzymes, are limited to the protein of interest and, once discerned, provide a powerful approach for purification of the protein. Other receptor or ligand sites, such as those with the ability to interact with certain dyes or lectins, respectively, are less specific, but nevertheless provide approaches for isolating particular proteins from complex mixtures of proteins. Receptor and/or ligand-binding sites present in a protein, such as reactivity with a particular substrate or monoclonal antibody, are often known prior to the initial biochemical characterization of the protein. This knowledge saves time and facilitates development of the purification procedure. In other cases, such as for dye binding or lectin reactivity, it is often necessary to perform pilot studies to determine the feasibility of using a particular affinity chromatography method. Because each protein contains multiple receptor and/or ligand-binding sites, the combination of several different affinity chromatography steps in a purification scheme can often be utilized to obtain pure proteins.

Receptor and ligand-binding sites are usually natural properties of proteins; however, it has become popular in the case of recombinant proteins to engineer particular binding sites, such as metal-binding sites, as extensions from either the N- or C-terminus into the recombinant protein of interest. Such alterations of the natural protein greatly facilitate their purification. If structural alteration of the recombinant protein by the ligand-binding site is undesirable, protease cleavage sites can be positioned adjacent to the ligand-binding site so that this protein extension can be removed after it has been used for affinity purification (e.g., as when a polyhistidine Ni^{2+} binding site is removed by thrombin cleavage).

Essentially all forms of affinity chromatography involve the following steps:

1. Choice of appropriate ligand.
2. Immobilization of the ligand onto a support matrix.
3. Loading of the protein mixture of interest onto the matrix.
4. Removal of nonspecifically bound proteins.
5. Elution of the protein of interest in a purified form.

This chapter presents an eclectic collection of affinity purification protocols that have proven extremely useful in protein purification.

BOJANA / CENTRAL UNIVERSITY LIBRARY

Contributed by John E. Coligan
Current Protocols in Protein Science (1998) 9.0.1-9.0.3
Copyright © 1998 by John Wiley & Sons, Inc.

Lectin affinity chromatography, described in *UNIT 9.1*, is frequently utilized in glycoprotein purification schemes. In this approach a protein is bound to an immobilized lectin through its sugar chain, unbound protein is washed away, and the bound protein is eluted. Sugar side chains present in glycoproteins are usually of sufficient complexity that they present the possibility of reacting with a number of different lectins in a fashion useful for purification. Suitable lectins are tabulated in this unit. However, the potential resolving power of this large battery of lectins is limited by the facts that (1) most or many glycoproteins in a crude mixture of proteins share identical or near-identical sugar side chains and (2) lectin specificities are often not absolute. Nonetheless, when appropriately applied, lectin affinity purification usually yields a significant enrichment of the glycoprotein of interest. The basic and alternate protocols describe the use of concanavalin A (Con A) or wheat germ agglutinin (WGA) for preparative glycoprotein purification. These lectins are commonly used because of their general applicability and availability. Others can be readily substituted using the procedures described.

Dye affinity chromatography, described in *UNIT 9.2*, is a protein purification procedure based on the high affinity of immobilized dyes for binding sites on many proteins. It is a rapid and versatile method that is especially applicable for obtaining 10- to 100-fold purification of the desired protein from crude extracts. The dye ligand can be selected either to remove unwanted proteins (negative chromatography) or to bind to the protein of interest (positive chromatography); these two types of chromatography can also be coupled in series (tandem chromatography). Support protocols are provided to guide selection of appropriate dyes and support matrices.

Affinity purification of proteins can also be accomplished by their interaction with natural ligands, as described in *UNIT 9.3*. These ligands may be macromolecules such as proteins and nucleic acids, or low-molecular-weight compounds such as substrates, inhibitors, cofactors, etc. Because such interactions tend to be specific for particular proteins or groups of proteins, high degrees of purification can be obtained in a single step. Because of this specificity, the protocols in *UNIT 9.3* do not detail a specific purification scheme, but deal instead with choices of column matrices and methods of coupling potential ligands to these matrices.

UNIT 9.4 describes a method for the facile purification of recombinant proteins by metal-chelate affinity chromatography (MCAC). Proteins engineered to have six consecutive histidine residues on either the amino or carboxyl terminus can be purified in native or denatured forms using a resin containing nickel ions (Ni^{2+}). Techniques for creating a fusion protein consisting of the protein of interest with a histidine tail are outlined in a Strategic Planning section. The protocols describe expression of the histidine-tailed fusion proteins and their purification in native form by MCAC, as well as the purification of the fusion protein under denaturing conditions and the subsequent renaturation procedure (by either dialysis or solid-phase renaturation). In addition, analysis of the purified product and regeneration of the resin is presented.

UNIT 9.5 describes the isolation of soluble or membrane-bound protein antigens by immunoaffinity chromatography. By this technique, purifications of 10,000-fold or more can often be achieved in one step. Immunoaffinity columns most often employ monoclonal antibodies covalently attached to a solid-phase matrix; polyclonal antibodies have been used successfully, but they usually lack the specificity required for a single-step protein purification. The amount of protein that can be purified is solely dependent on the amount and affinity of the antibody employed. In addition to describing the overall procedure, this unit provides alternate conditions of elution and a method for attaching the antibody to the support matrix.

BCU IASRI CENTRAL UNIVERSITY LIBRARY

Many biological processes, such as recombination, replication, and transcription, involve the action of sequence-specific DNA-binding proteins. Affinity chromatography (as described in *UNIT 9.6*) is a very effective means of purifying a protein based on its sequence-specific DNA-binding properties, and is relatively straightforward if proper care is taken. The affinity chromatography procedure that is described in this unit uses DNA encoding specific recognition sites for the desired protein that has been covalently linked to a solid support. Researchers in the past have purified proteins based on the proteins' ability to bind DNA, only to be disappointed to find that the DNA-binding activity of their highly purified products was not sequence-specific. This situation can be avoided by carefully performing all of the preliminary steps and experimental controls described in this unit. Figure 9.6.1 provides a summary diagram of the entire chromatography procedure. With this procedure, a high degree of purification can be achieved using a DNA affinity column.

A similar degree of purification can be accomplished using biotinylated DNA fragments that contain the desired protein binding site and streptavidin-coated column matrices (*UNIT 9.7*). This purification scheme, outlined in Figure 9.7.1, utilizes the tight and nearly irreversible binding that occurs between biotin and streptavidin to fish out specific DNA binding proteins from crude cellular or nuclear lysates. After washing, the protein of interest is eluted with high-salt buffer. This protocol is relatively easy to optimize by following the success of the isolation using gel mobility-shift assays; it is also economical, as a single type of column matrix is compatible with any biotinylated DNA probe.

In summary, affinity chromatography at its best is the most powerful technique for protein purification available, because its high selectivity can, in principle, permit purification of a desired protein from a crude mixture of proteins in a single step. Moreover, the technique usually offers simultaneous concentration of the desired protein.

John E. Coligan

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Lectin Affinity Chromatography

Lectin affinity chromatography is the most frequently used specific purification procedure for glycoproteins. A protein is bound to an immobilized lectin through its sugar chain, the unbound protein is washed away, and the bound protein is eluted. Numerous plant lectins have been identified and in many cases their sugar specificity is known. This sugar specificity is the basis for the separation and structural analysis of many individual oligosaccharides and glycopeptides (Chapter 12) and for the identification of lectin-binding glycoproteins. Glycoproteins often contain multiple chains of a given sugar type; the resulting multiple interactions with the immobilized lectin may make it difficult to elute some of the proteins from the lectin column. The presence of multiple low-affinity sugar chains can still produce substantial interaction with a lectin-containing column even though individual sugar chains may have a low affinity. Thus, only limited information on the structure of a sugar chain can be obtained from the binding of a protein to a lectin column. These qualifications by no means belittle the usefulness of lectin affinity chromatography for glycoprotein purification; lectin affinity chromatography may be used to partially purify glycoproteins and to provide limited qualitative information about the nature of their carbohydrate components. Other preparative purification steps, such as ammonium salt precipitation and gel-filtration and ion-exchange chromatographies (*UNIT 8.2 & UNIT 8.3*), are usually done prior to using lectins, but this is not necessary.

Many different lectins are commercially available in an immobilized form suitable for glycoprotein purification, but some can be rather expensive. Con A–Sepharose is the most commonly used lectin for glycoprotein purification. It is relatively inexpensive, stable, and able to bind to many different glycans. Bound proteins can be eluted with α-methyl-D-mannoside (αMM). Wheat germ agglutinin (WGA) is the next most popular lectin for the same reasons although it is somewhat more expensive.

Many other lectins can be used to purify glycoproteins. A broad range of immobilized lectins is available from E-Y Labs, Vector, Sigma, and Pharmacia Biotech (*SUPPLIERS APPENDIX*). Most are also available in free form, if desired. However, this array of lectins is never used for routine protein purification. First, their higher price, shorter track records, and, in some cases, requirement for exotic sugars to elute bound proteins make them less popular for routine use. Second, the time and cost of screening and optimizing conditions for so many lectins would be prohibitive. Third, many have very similar specificities. Fourth, even using a battery of lectins to purify a protein would not yield a single species, since other proteins with similar sugar chains would copurify. Although these other lectins are less frequently used, the principles of Con A–Sepharose and WGA–agarose lectin affinity chromatography apply to them as well.

The names, sugar specificity, elution conditions, and general types of sugar chains that bind to 25 different lectins are given in Table 9.1.1. This table is not comprehensive (there are at least twice as many lectins commercially available), but those listed are the most commonly used. The list is intended as a guide to possible alternatives, and an attempt has been made to group lectins of similar specificities together. However, specificities are not always hard and fast, and the comments in the table are not always sacrosanct. Some lectin–ligand binding requirements are well characterized, but many of the "rules" for lectin sugar specificity are based on screening simple sugars to inhibit lectin binding. These competitors can mimic only limited features of the binding of a lectin to its natural ligand. For instance, both pea and lentil lectins bind only sugar chains with α-fucose residues in the core region of N-linked chains, and yet mannose, not fucose, is used to elute bound proteins. Cummings (1994) provides an excellent perspective on the use of lectins.

Copyright © 1995 by John Wiley & Sons, Inc.

BUSWELL CENTRE UNIVERSITY LIBRARY

Table 9.1.1 Lectins Useful for Glycoprotein Purification

Acronym (organism and source)	Metal ions required	Sugar specificity[b]	Elution conditions	Useful for binding
Con A (*Canavalia ensiformis;* jack bean seeds)	Ca^{2+}, Mn^{2+}	α-Man > α-Glc	0.1–0.5 M α-MeMan	High-Man, hybrid, and biantennary N-linked chains
LCA or **LCH** (*Lens culinarus;* lentil seeds)	Ca^{2+}, Mn^{2+}	α-Man > α-Glc	0.1–0.5 M α-MeMan	Bi- and triantennary N-linked chains with Fucα1-6 in core region
PSA (*Pisum sativum;* peas)	Ca^{2+}, Mn^{2+}	α-Man	0.1–0.5 M α-MeMan	Similar to LCA/LCH
WGA (*Triticum vulgaris;* wheat germ)	Ca^{2+}, Mn^{2+}	β-GlcNAc	0.1–0.5 M GlcNAc	GlcNAc- and Sia-terminated chains, or clusters of *O*-GlcNAc; succinylated form selectively binds GlcNAc>Sia
RCA (*Ricinus communis;* castor bean)				
RCA-I[a]	—	β-Gal β-GalNAc	0.1–0.5 M Gal or lactose	Galβ1-4 on N-linked chains; can also bind β-GalNAc
RCA-II[a]	—	β-Gal	0.1–0.5 M Gal	β-GalNAc-terminated chains.
PHA (*Phaseolus vulgaris;* red kidney bean)				
E_4PHA	—	GalNAc	0.1–0.5 M GalNAc	Gal-terminated N-linked chains with "bisecting" GlcNAcβ1-4
L_4PHA	—	GalNAc	0.1–0.5 M GalNAc	Tri- and tetrabranched N-linked chains
GSL-1, BSL-1 [*Griffonia (Bandeiraea);* Simplicifolia seeds]	Mg^{2+}, Ca^{2+}	α-Gal α-GalNac	0.1–0.5 M melibiose	Chains terminated by α-Gal residues
DSA (*Datura stramonium;* Jimsonweed seeds)	—	Galβ1-4GlcNAc	0.1–0.5 M GlcNAc or 20 mM *N,N′,N″*-chitotriose	Proteins with poly-*N*-acetyllactosamine-containing chains; highly heterodispersed glycoproteins often contain variable lengths of this repeat
LEL (*Lycopersicon esculentum;* tomato meat)	—	β-GlcNAc	0.1–0.5 M β-GlcNAc or 20 mM chitobiose or chitotriose	Similar to DSA; prefers longer poly-*N*-acetyllactosamines than DSA
STL (*Solanum tuberosum;* potato tubers)	—	β-GlcNAc	0.1–0.5 M GlcNAc or 20 mM chitobiose or chitotriose	Proteins with poly-*N*-acetyllactosamines
MAL (*Maackia amurensis;* seeds)	—	Siaα2-3Galβ1-4 GlcNAc	0.5 M lactose	α2-3-linked Sia to Gal in N-linked sugar chains; underlying sugars and peptide contribute to binding

continued

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 9.1.1 Lectins Useful for Glycoprotein Purification, continued

Acronym (organism and source)	Metal ions required	Sugar specificity[b]	Elution conditions	Useful for binding
EBL or **SNA** (*Sambucus nigra;* elderberry bark)	—	Siaα2-6Gal or GalNAc	0.1–0.5 lactose	α2-6-linked Sia in N and O sugar chains
LFA (*Limax flavus;* garden slug)	—	Neu5Ac	10 mM Sia	Sia-terminated chains irrespective of linkage
GNL (*Galanthus nivalis;* snowdrop bulbs)	—	α1-3Man	0.5 M α-MeMan	High-Man-type chains, but not to Glc
UEA-I (*Ulex europaeus;* furze gorse seeds)	—	α-L-Fuc	0.1–0.5 M L-Fuc or methyl-α-L-Fuc	Sugar chains with terminal α-Fuc, especially in α1-2 linkage, but much less with α1-3 or α1-6 linkages
AAA (*Anguilla anguilla;* freshwater eel)	—	α-L-Fuc	0.1–0.5 M L-Fuc	Similar to UEA-I but more broadly binds any fucosylated oligosaccharide
Lotus (*Lotus tetragonolobus;* asparagus pea)	—	α-L-Fuc	0.1–0.5 M L-Fuc	Outer-branch α-Fuc residues; does not bind when Fuc is linked to chitobiose core of N-linked chains
HPA (*Helix promatia;* albumin gland of edible snail)	—	α-GalNAc	0.1–0.5 M GalNAc	Proteins with terminal α-GalNAc or GalNAcα-*O*-Ser/Thr (Tn antigen)
Jackalin (*Artocarpus integrifolia;* jackfruit seeds)	—	α-Gal	0.1–0.2 M melibiose	Proteins with only a single O-linked chain with T antigen Galβ1-3GalNAcα-*O*-Ser/Thr can bind; not inhibited by lactose; Sia in chain does not interfere with binding
VVL or **VVA** (*Vicia villosa;* hairy vetch seeds)	—	GalNAc	0.1–0.5 M GalNAc	Proteins with Tn antigen, GalNAcα-*O*-Ser/Thr
SBA (*Glycine max;* soybean)	—	GalNAc or Gal	0.1–0.5 M GalNAc	α- or β-linked GalNAc
PNA (*Arachis hypogaea;* peanut)	—	Galβ1-3GalNAc	0.1–0.5 M lactose	Proteins with T antigen, but does not bind if sugar chain is sialylated
DBA (*Dolicholos biflorus;* horse gram seeds)	—	Terminal α-GalNAc	0.1–0.5 M GalNAc	Terminal α-GalNAc (specific for blood group A), not those in the core of O-linked chains
LBA (*Phaseolus lunatus;* lima bean)	Mn^{2+}, Ca^{2+}	Terminal α-GalNAc	0.1–0.5 M GalNAc	Proteins with blood group A structure GalNAcα1-3(Fucα1-2)Gal–

[a]*CAUTION:* Both RCA-I and RCA-II are extremely toxic, with lethal doses of 1 molecule/cell.

[b]Abbreviations: Fuc, L-fucos; Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; Glc, D-glucose; GlcNAc, *N*-acetyl-D-glucosamine; Man, D-mannose; Neu5Ac, *N*-acetyl-D-neuraminic acid; Sia, sialic acid.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

In general, these immobilized lectins can be used in two ways. First is as an affinity step for routine purification. If a novel protein may be glycosylated, using several lectins that bind very different types of sugar structures (N-linked vs. O-linked) will probably give reasonable purification. It will also give some indication of the type of sugar chains on the protein. Both Vector and E-Y Labs sell kits with six or seven different lectins for this purpose. The second and more systematic approach is useful for better-characterized glycoproteins. Here, the lectins are chosen to resolve potentially different glycoforms that may have different biological activities. An excellent example of this second type of analysis is given by Li and Jourdian (1991).

The Basic and Alternate Protocols describe the use of lectins for preparative glycoprotein purification. Con A–Sepharose and WGA–agarose were chosen for convenience and availability. The Support Protocol describes a small-scale pilot procedure to test for lectin binding and to determine elution conditions. There are many variations on the basic procedure in the literature, but all use the same principles: bind the protein to immobilized lectin through its sugar chain, wash away unbound protein, and elute bound protein with a simple sugar that resembles the sugar ligand of the bound protein. Because many proteins have sugar chains that can bind to a specific lectin, this procedure seldom yields a pure protein.

BASIC PROTOCOL

CON A–SEPHAROSE AFFINITY CHROMATOGRAPHY

Con A–Sepharose chromatography is used to partially purify glycoproteins that contain terminal mannose or glucose residues. The steps presented below give typical conditions for this type of chromatography, although a variety of approaches have been utilized. In this protocol, bound glycoproteins are eluted with α-methyl-D-mannoside (αMM) after the column is first washed to remove unbound and weakly bound proteins. Before proceeding, it is advisable to conduct a pilot study to test the protein of interest for lectin binding and elution conditions (see Support Protocol). In order to purify the protein effectively, a specific assay for detecting the protein of interest (i.e., activity or antibody-binding assay) must be available.

Materials

10 mg/ml Con A–Sepharose (Pharmacia Biotech or Sigma)
Column buffer (see recipe)
0.5 M αMM in column buffer
Protein sample in column buffer

1.5 × 30–cm glass or disposable chromatographic column
Glass wool

NOTE: Carry out this procedure at room temperature if the protein to be isolated will tolerate this condition. If not, carry it out in a cold room, and prechill all solutions to maintain temperature.

1. Gently resuspend 50 ml settled Con A–Sepharose (10 mg lectin/ml packed resin) in 50 ml column buffer to make a slurry. Degas the slurry.

   This volume of lectin beads should be sufficient to bind ~100 mg glycoprotein. If a 1.5 × 30–cm column is not available, either the dimensions of the column or the total amount of resin can be modified. A 1.0 × 30–cm column will bind ~50 mg total glycoprotein. If the amount of glycoprotein in the sample is substantially less than 50 to 100 mg, decrease the volume of the column accordingly. The ratio of input protein to lectin does not seem to matter as long as the column is not overloaded.

BCU IASTOCENTRAL UNIVERSITY LIBRARY

2. Pack a glass wool plug over a sintered glass or polypropylene frit at the bottom of the column and pour the degassed slurry into the column.

   The glass wool prevents the beads from clogging the frit and slowing the flow rate. Alternatively, BioGel P2 or Sephadex G10 can be used in place of the glass wool.

3. Continue packing the column until the desired level is reached (for a 50-ml volume this is ~28 cm). Wash the gel with 2 to 3 column volumes of column buffer to remove any loosely bound or degraded Con A.

4. Wash with 2 to 3 column volumes of 0.5 M αMM in column buffer or the highest concentration of αMM that will be used.

5. Wash the column with >5 column volumes of column buffer *without* αMM to reequilibrate.

   It is important to prewash the column with the eluting sugar and then to reequilibrate the column to remove any materials that may have previously bound to the column. Check completeness of washing by phenol–sulfuric acid assay (Manzi, 1993). Estimate the amount of residual sugar using αMM as a standard. An acceptable level is <0.1 mM or ~20 μg/ml.

6. Slowly load the protein sample on the column to permit binding without disturbing the surface.

   A flow rate of ~1 ml/min (~0.5 ml/min/cm^2 of area) is desirable.

7. Wash the column with column buffer and monitor the flowthrough and subsequent wash fractions by measuring the A_{280} until it approaches baseline value.

   If the A_{280} does not rapidly return to baseline, it may indicate weak interaction of some proteins with the lectin, or overloading of the column.

8. Using an appropriate specific assay, check flowthrough and wash fractions for the presence of the protein of interest (e.g., enzyme activity or antibody binding).

   To check the lectin-binding properties, see Support Protocol. To determine whether the column was overloaded, add a small amount of the flowthrough (1%) to a small volume of beads as described in the Support Protocol. If the sample is bound to the beads as determined by the appropriate assay, overloading of the original column is indicated. In this case, reapply the flowthrough with excess glycoprotein to the larger column after the bound proteins of the first run have been eluted with αMM and the column reequilibrated.

9. Elute the column with 0.5 M αMM in column buffer. Monitor fractions for A_{280} and for activity (using specific assay) and pool peak-activity fractions.

   Very broad peaks may result during elution with the sugar if the protein dissociates very slowly. When this happens, sample recovery can be improved by filling the column with 0.5 M αMM in column buffer and allowing the column to stand for a few hours. This allows dissociation of the bound material; when the flow is started again, it should give a very sharp peak. Other possible remedies include warming a cold column to room temperature in the presence of αMM, increasing the concentration of αMM to 1 M, or increasing the NaCl concentration to 1 M. A combination of these approaches may be needed to elute a tightly bound protein.

10. Regenerate the column by washing it with 10 vol column buffer or until the αMM concentration is <20 μg/ml.

    The column may now be used for another run or stored indefinitely at 4°C in column buffer containing 0.02% (w/v) NaN_3. Reequilibrate in column buffer before using again.

BOWLING GREEN STATE UNIVERSITY LIBRARY

PILOT STUDY TO DETERMINE LECTIN BINDING AND ELUTION CONDITIONS

If the target protein can be detected easily, test a small sample to establish binding and elution conditions before applying the entire sample to a large column. The easiest approach is to mix a small amount of sample with a measured amount of Con A–Sepharose beads in a series of microcentrifuge tubes. If the material in the sample binds, elute it from the washed beads with various amounts of competing αMM to determine when the activated material is eluted into the supernatant. This shows whether binding occurs and what concentration of competing sugar (or other conditions) will be required to elute it from the column. An abbreviated version of this procedure can be done with a single sample to determine if the protein binds and if it can be eluted with only one concentration (0.75 M) of αMM. It is important to keep the temperature constant, because changes can affect binding and dissociation of the ligands.

Additional Materials (also see Basic Protocol)

Sepharose 4B (Pharmacia Biotech) or other beaded gel to fill space in the tubes
αMM: 0, 0.1, 0.2, 0.4, 0.8, 1.0, and 1.5 M concentrations, in column buffer

1. Prepare a 50% slurry of Con A–Sepharose in column buffer. Cut ~2 to 3 mm off the small end of a 1000-µl pipet tip. Resuspend the slurry immediately before each pipetting, and use the truncated tip to dispense 200 µl of slurry into a series of seven microcentrifuge tubes. Add an equal volume of 50% slurry of Sepharose 4B to one tube (to serve as no-lectin control). Allow gel to settle to see if the dispensing was reasonably accurate.

   Gel material is needed in the controls to account for the volume changes caused by adding the gel. Other nonionic beaded gels, such as Sephadex (Pharmacia Biotech) or Bio-Gel (Bio-Rad) may be substituted.

2. Add the protein sample in 50 µl column buffer and allow 15 min for binding (shake the tube occasionally). Microcentrifuge the gel beads 1 min at 1000 × *g* to sediment. Remove the supernatant and save for analysis (step 6).

3. Use an appropriate assay specific for the protein of interest to determine if it binds to Con A–Sepharose and not to control beads.

4. Wash gel beads by adding 1.4 ml column buffer, microcentrifuging 1 min at 1000 × *g*, and removing supernatant (save for analysis). Repeat wash twice more.

5. Resuspend gel beads in 200 µl of buffer containing 0, 0.1, 0.2, 0.4, 0.8, 1.0, or 1.5 M αMM for each of the seven tubes containing Con A–Sepharose. Add the same volume of column buffer or 1.5 M αMM in column buffer to the control.

6. Incubate 15 min and centrifuge the gel beads 1 min at 1000 × *g*. Remove supernatant. Analyze the supernatant (along with those from steps 2 and 4) using the appropriate detection assay to determine if the protein has been eluted and at what concentration of competing sugar this occurred.

   Increasing the concentration of αMM should elute the target protein from the Con A–Sepharose. If it is not eluted, longer incubation (10 hr) in the presence of αMM or higher temperature may be needed. Although the entire procedure can be done in the cold, binding is tighter at low temperature, and this may make elution more difficult. Once the conditions are established for a particular sample or a lectin, scaling up the preparation for the basic protocol is normally routine.

BCU IASI CENTRAL UNIVERSITY LIBRARY

WHEAT GERM AGGLUTININ (WGA)–AGAROSE AFFINITY CHROMATOGRAPHY

ALTERNATE PROTOCOL

WGA-agarose chromatography is used to purify proteins that contain terminal *N*-acetyl-D-glucosamine (GlcNAc) or sialic acid residues. A protein sample is applied to the column, the column is washed to remove unbound and weakly bound proteins, and bound glycoproteins are eluted with GlcNAc. The advice for purification of glycoproteins on Con A–Sepharose columns applies to columns containing immobilized WGA. Test the binding of the target protein to the lectin using a pilot study like that for Con A–Sepharose (see Support Protocol) before running a large column, substituting GlcNAc for the αMM.

Additional Materials (also see Basic Protocol)

5 mg/ml wheat germ agglutinin (WGA)–agarose (E-Y Labs, Pharmacia Biotech, or Sigma)
PBS (*APPENDIX 2E*)
0.1 M *N*-acetyl-D-glucosamine (GlcNAc) in PBS
Protein sample in PBS
1.0 × 10–cm glass or disposable column (or other column whose length is 10 times its diameter)

1. Resuspend 1 vol WGA–agarose in 2 to 3 vol PBS to make a slurry.

   One milliliter of gel should be enough to bind 1 to 2 mg glycoprotein. To be safe, assume ~5% to 10% of total cell lysate can bind. If target protein requires detergent to solubilize it, WGA is still active in 25 mM Tris·Cl (pH 7.5)/1% Lubrol PX/0.1% sodium deoxycholate. Triton X-100 or Nonidet P-40 (~2% to 3% v/v final) may also be used.

2. Place a glass wool plug over the frit at the bottom of a column whose length is 10 times its diameter. Pour the slurry into the column. Wash the column with ~2 column volumes of PBS to remove any unbound or degraded WGA, then wash with 1 column volume PBS containing 0.1 M GlcNAc. Finally, wash with 5 column volumes PBS.

   The length of the column should be ~10 times its diameter so that weakly bound proteins can be eluted by prolonged PBS wash in the absence of any competing sugar. This will separate them from proteins that do not interact, and differs from using Con A–Sepharose, where continued buffer wash in the absence of αMM does not usually elute weakly bound protein. The choice of these column dimensions for WGA and for most other lectins is based on the successful fractionation of individual glycopeptides by serial lectin affinity chromatography.

   A plug of BioGel P2 or Sephadex G10 can be used in place of the glass wool.

3. Add protein sample in ~0.1 column volumes at a flow rate of ~2 ml/cm^2 per hour, and wash the column with PBS until the A_{280} returns to baseline.

   A small sample volume makes it easier to detect proteins that do not bind to the column and elute promptly as a sharp peak in the flowthrough. Proteins that bind weakly to the column are gradually washed off by continued elution with PBS alone; these proteins would "smear" as a trailing peak. If a preliminary test run shows that the target protein binds strongly to the lectin, and if the column is not overloaded, the sample can be added to the column in a larger volume. The flow rate can be varied. Sometimes the protein of interest can be absorbed to the lectin by gently mixing the two components for several hours before pouring the loaded gel into the column. The protein can then be eluted after washing as determined by the pilot study.

4. Elute the column with 2 to 3 column volumes of 0.1 M GlcNAc in PBS, and assay the fractions for the protein of interest.

   Generally, 0.1 M GlcNAc is sufficient to elute the protein. If this does not work, try a higher concentration (e.g., up to 0.25 M GlcNAc), temporarily turn off the column, or

BOSTON COLLEGE UNIVERSITY LIBRARY

raise the temperature to elute the sample, as with Con A–Sepharose. Adding 0.5 M NaCl to the GlcNAc-containing elution buffer can sometimes improve recovery. A pilot study to determine elution conditions, similar to that described for Con A–Sepharose (see Support Protocol), is highly recommended.

5. Regenerate the column by washing with ≥10 column volumes of PBS or until reducing sugar content is below 20 µg/ml. The column can be reused immediately or stored at 4°C in column buffer containing 0.02% (w/v) NaN_3.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Column buffer

0.01 M Tris·Cl, pH 7.5
0.15 M NaCl
1 mM $CaCl_2$
1 mM $MnCl_2$
Store indefinitely at room temperature

Prepare $MnCl_2$ within a day or two of use and add it to the buffer only after the pH has been adjusted. The final concentration of metal ions can be decreased 10-fold or more if needed.

If detergents are required to solubilize the protein, either Triton X-100 or Nonidet P-40 at ≤2% has a negligible effect on lectins. Do not use glucoside-based detergents because they may interfere with binding.

COMMENTARY

Background Information

See the introductions to *UNIT 12.1* and this unit for discussions concerning properties of glycoproteins that relate to their purification.

Critical Parameters

As with most protein purification procedures, successful lectin affinity chromotography is empirical, so few parameters are written in stone. However, in using lectins, it is important to first determine the binding and elution conditions in a pilot study. Blindly loading a sample onto a column and then trying different elution conditions may give very low yields, either because the protein becomes inactivated or because it cannot be eluted under those particular conditions. It is usually easy to bind the protein to the column, but elution may be more difficult. There are numerous ways to elute tightly bound protein: increase the concentration of sugar in the buffer used for elution, increase temperature, increase NaCl concentration, or stop the column flow for an extended period during the elution with sugar. A combination of these approaches may be needed.

These modifications may cause a protein to elute differently from the same column. Thus, it may be useful to perform two consecutive runs with different elution conditions to separate the protein of interest from different contaminants. Ketcham and Kornfeld (1992) describe the purification of a rare glycosyltransferase using WGA, whereby the lectin affinity chromatography step yielded a 160-fold purification. In addition, environmental carbohydrate contamination may contribute significantly to the total quantity of carbohydrate in a preparation. Follow the suggestions given within the protocols to minimize this effect.

A few general points to remember. Most lectin–sugar binding relies on hydrophobic interactions, so moderate (0.2 M) to high (0.6 M) salt concentrations do not prevent binding. A high salt concentration also reduces nonspecific protein binding. Similarly, tightly bound proteins may be eluted by including 50% ethylene glycol with the sugar. The strength of lectin binding can be significantly affected by the density of lectins on the beads. Usually 3 to 5 mg/ml will bind most proteins, but binding individual free sugar chains may require a higher concentration. Most useful lectins are supplied already bound to beads. If you must couple your own, CNBr-activated Sepharose 4B or Affi-Gel 10 (Bio-Rad) are good supports. Be sure to include any required metal ion and 0.2 M of the eluting sugar during the coupling.

BCU IAS / CENTRAL UNIVERSITY LIBRARY

This will help generate the active conformation and prevent coupling the lectin through the sugar-binding site. Also, keep the soluble lectin concentration below 10 mg/ml during coupling or some may precipitate. Finally, do not couple free *Ricinus communis* II (RCA-II) from castor beans in your own lab because it is deadly: the LD_{50} in mice is 0.8 μg/100 g body weight. For all lectin column runs avoid high concentrations (>1%) of nonionic detergents or any SDS.

Anticipated Results

Lectins can give 2- to >100-fold purifications of glycoproteins and nearly quantitative yields, depending upon the source of the sample and the elution conditions. If the protein of interest is reasonably stable at high salt concentrations and room temperature, specific elution conditions can likely be found that will produce good yields.

Time Considerations

Initial determination of the binding and elution conditions can easily be done in 1 day, assuming the detection assay for the target protein is simple and fast. The amount of time needed to conduct a single run once the elution conditions are known can vary depending upon the size of the sample and the column, and whether elution requires long-term incubation in the presence of the sugar. Even assuming the extremes of sample size and elution conditions, a single run should be completed in 2 to 3 days.

Literature Cited

Cummings, R.D. 1994. Use of lectins in analysis of glycoconjugates. *Methods Enzymol.* 230:66-86.

Ketcham, C.M. and Kornfeld, S. 1992. Purification of UDP-*N*-acetylglucosamine: Glycoprotein *N*-acetylglucosamine-1-phosphotransferase from *Acanthamoeba castellanii* and identification of a subunit of the enzyme. *J. Biol. Chem.* 267:11645-11653.

Li, M. and Jourdian, G.W. 1991. Isolation and characterization of the two glycosylation isoforms of low molecular weight mannose-6-phosphate receptor from bovine testis. *J. Biol. Chem.* 266:17621-17630.

Manzi, A.E. 1993. Phenol–sulfuric acid assay for hexoses and pentoses. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 17.9.1-17.9.2. John Wiley & Sons, New York.

Key References

Beeley, J.G. 1984. Glycoprotein and proteoglycan techniques. *In* Laboratory Techniques in Biochemistry and Molecular Biology, pp. 29-99. Elsevier Science Publishing, New York.

Good discussions about general properties of glycoconjugates.

Cummings, R.D. 1994. See above.

Most current description of using lectins to characterize sugar chain structures. Lots of details on procedures to analyze free sugar chains rather than glycoproteins.

Dulaney, J.T. 1979. Binding interactions of glycoproteins with lectins. *Mol. Cell. Biochem.* 21:43-62.

Lists many references and conditions used for lectins in protein purification.

Li, M. and Jourdian, G.W. 1991. See above.

Excellent example of separation of various glycoforms using lectin–affinity chromatography.

Montreuil, J., Bouquelet, S., Debary, H., Fournat, B., Spik, G., and Strecker, G. 1986. Glycoproteins. *In* Carbohydrate Analysis: A Practical Approach (M.F. Chaplin and J.F. Kennedy, eds.) pp. 166-173. IRL Press, Washington, D.C.

Lists several conditions for selected lectin-affinity purifications of proteins.

Contributed by Hudson H. Freeze
La Jolla Cancer Research Foundation
La Jolla, California

B.C.U. „M. ... SCU" IAŞI
CHI... IE

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Dye Affinity Chromatography

Dye affinity chromatography is a protein purification procedure based on the high affinity of immobilized dyes for the binding sites on many proteins. It is a rapid, inexpensive, and versatile method that is applicable to the purification of crude cellular extracts.

There are three types of dye affinity chromatography—negative chromatography, positive chromatography, and tandem chromatography. Negative chromatography (Basic Protocol 2) is the simplest of the three procedures. Some of the undesired proteins are retained by the immobilized dye while the remaining proteins, including the desired protein, flow through the column. This results in a modest-fold purification. Negative chromatography is particularly convenient for the rapid removal of degradative enzymes such as proteases and nucleases, and for the removal of very abundant proteins, such as serum albumin. A different immobilized dye is used for positive chromatography (Basic Protocol 3 and Alternate Protocol 2). This immobilized dye retains the desired protein as well as some undesired proteins. The desired protein is then selectively eluted using either a nonspecific reagent, such as a neutral salt, or a specific reagent, such as a substrate, cofactor, or competitive inhibitor. Although positive chromatography requires more investigator involvement, it results in a greater-fold purification than obtained using negative chromatography, and is more commonly used. In some cases, a desired protein can be purified to homogeneity from a crude cellular extract using positive chromatography.

In tandem chromatography (Basic Protocol 4), the most elaborate of the three procedures, negative chromatography and positive chromatography are coupled sequentially. The protein mixture first passes through a column containing the immobilized dye chosen for negative chromatography. The desired protein flows through this column and immediately enters another column containing the immobilized dye chosen for positive chromatography. The desired protein, which is retained by the latter column, is then eluted using either batch or gradient elution. Tandem chromatography is particularly useful for situations in which the immobilized dye used for positive chromatography binds too many undesired proteins, making selective elution of the desired protein difficult.

Use of dye affinity chromatography for protein purification requires selection of one or more immobilized dyes. In addition, positive chromatography requires selection of an elution reagent and its effective concentration. These selections can be made using chromatography, as outlined in Basic Protocol 1, or using centrifugation, as outlined in Alternate Protocol 1, depending on the inclination of the investigator. Immobilized dyes can be purchased as dry powders or suspensions or already poured into miniature columns in kit form. Free dyes also can be purchased and easily immobilized by the investigator (Support Protocol). Alternate Protocol 1 facilitates selection because many samples can be processed simultaneously, but is more expensive because more immobilized dye is required.

SELECTION OF COMPONENTS USING CHROMATOGRAPHY

BASIC PROTOCOL 1

The immobilized dye for negative or positive chromatography and the eluant for positive chromatography can be selected using a chromatographic procedure. A small volume of the protein mixture to be purified is applied to a series of miniature columns, each containing a different immobilized dye. The flowthrough and washes from each column are pooled and analyzed for total protein and for desired protein. The immobilized dye column whose pooled flowthrough and washes contain the least amount of undesired protein and the greatest amount of desired protein is appropriate for negative chromatography. Conversely, the immobilized dye column whose pooled liquid contains the greatest

Copyright © 1995 by John Wiley & Sons, Inc.

BYU/CENTRAL UNIVERSITY LIBRARY

amount of undesired protein and the least amount of desired protein is appropriate for positive chromatography.

The immobilized dye column chosen for positive chromatography is used to select the elution reagent and its effective concentration. The column is washed with small volumes of a potential elution reagent, with the concentration of the reagent increased in small increments in each successive wash. The flowthrough for each wash is then analyzed for both total protein and desired protein. Additional potential elution reagents are examined following application of a fresh aliquot of the protein mixture to be purified. The concentration of elution reagent that provides the greatest-fold purification, recovery of desired protein, and economy of use is selected to be the elution solvent. A specific elution reagent will generally provide a greater increase in fold purification than a nonspecific elution reagent, but the specific elution reagent may cost more.

Materials

Protein mixture to be purified
Appropriate application solvent (see step 1)
Series of immobilized dyes (Table 9.2.1 and Support Protocol)
Concentrated solution of an appropriate elution reagent (see step 11)
Plastic disposable filtration unit with pore size ≤0.45 μm
Small plastic disposable chromatography tubes with bed volume at least 1 ml

Table 9.2.1 Some Commercially Available Immobilized Dyes[a]

Dye	Immobilized dye	Supplier[b]
Reactive blue 2	Affi-Gel blue gel	Bio-Rad
	Blue Sepharose CL-6B	Pharmacia Biotech
	Cibacron blue 3G-A-agarose	ICN Biomedicals, Sigma
	Cibacron blue F3G-A	Pierce
	DyeMatrex blue A gel	Amicon
Reactive blue 4	Reactive blue 4-agarose	Sigma
Reactive blue 72	Reactive blue 72-agarose	Sigma
Reactive brown 10	Reactive brown 10-agarose	Sigma
Reactive green 5	Reactive green 5-agarose	ICN Biomedicals, Sigma
Reactive green 19	DyeMatrex green A gel	Amicon
	Reactive green 19-agarose	Sigma
Reactive red 120	DyeMatrex red A gel	Amicon
	Reactive red 120-agarose	ICN Biomedicals, Sigma
	Red Sepharose CL-6B	Pharmacia Biotech
Reactive yellow 2	Reactive yellow 2-agarose	ICN Biomedicals, Sigma
	Reactive yellow 3-agarose	Sigma
Reactive yellow 3	DyeMatrex orange A gel	Amicon
Reactive yellow 13	Reactive yellow 13-agarose	Sigma
Reactive yellow 86	Reactive yellow 86-agarose	Sigma

[a]Screening kits containing between 5 and 40 immobilized dyes in miniature chromatography columns can be purchased from Affinity Chromatography, Amicon, Wako Pure Chemical, and Sigma. Alternatively, free reactive dyes such as those listed in Table 9.3.2 can be purchased, immobilized, and prepared for chromatography as described in the Support Protocol.

[b]Supplier names and addresses are provided in *SUPPLIERS APPENDIX*.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Prepare protein sample for application to column

1. Prepare the crude protein mixture. If it is a solid, dissolve it in the application solvent. If the protein mixture to be purified is a suspension or a solution, dialyze it for several hours against the application solvent.

   A generic application solvent is 50 mM Tris·Cl buffer, pH 7.5, at room temperature. The identity of the buffer, its pH, and the temperature of the application solvent may be varied to maintain the integrity of the desired protein without compromising the chromatography. However, addition of reagents that increase the ionic strength of the application solvent above 50 mM may diminish retention of all proteins by the immobilized dye columns. Addition of millimolar concentrations of one or more specific reagents that bind to the desired protein to maintain its integrity may diminish the retention of the desired protein by the immobilized dye columns.

2. Clarify the protein mixture, if necessary, either by centrifugation or filtration.

   Filtration should be performed using a plastic disposable filtration unit having a pore size of ≤0.45 μm.

3. Determine the amount of total protein and the amount of desired protein in 1 ml of the protein mixture.

Prepare assortment of immobilized dye columns

4. Pour a sufficient volume of each immobilized dye suspension into a different chromatography tube to form a bed volume of 1 ml in each tube.

   Most commercial screening kits of immobilized dyes are supplied as chromatographic columns having bed volumes ranging from 0.1 to 2.5 ml; these may be used without alteration. The optimal bed volume for a particular situation depends on the total volume of the crude protein mixture available to the investigator; the greater this total volume, the more that can be used. The volume of available crude protein mixture is the principal reason the bed volumes of individual columns in the commercial kits range from 0.1 to 2.5 ml.

   If one or more of the immobilized dyes are available in powdered form, suspend about 0.5 g of each powdered immobilized dye in application solvent for at least 30 min with gentle stirring prior to pouring the columns.

5. Wash each immobilized dye column with 2 bed volumes of chromatographic application solvent. Let the column run dry and discard the washes.

Perform chromatography and assess column performance

6. Apply the same volume of protein mixture to each immobilized dye column and collect the flowthrough in a container.

   The volume of protein mixture added should contain between 5 and 40 mg of total protein per milliliter of bed volume and sufficient desired protein (1% of total protein) for convenient analysis. The flowthrough can be collected in a disposable tube of sufficient volume to contain the application volume and 2 bed volumes of wash.

7. Wash each immobilized dye column with 2 bed volumes of application solvent and collect the flowthrough in the same container used in step 6.

8. Measure the concentrations of the total protein and the desired protein in the combined flowthrough from each column.

9. Compare the amount of total and desired protein in the combined flowthrough from each column.

   The immobilized dye whose combined flowthrough contains the least amount of total protein and the greatest amount of desired protein is most appropriate for negative chromatography (see Basic Protocol 2). The immobilized dye whose combined

BCU MSU ET TRAU UNIVERSITY LIBRARY

flowthrough contains the greatest amount of total protein and the least amount of desired protein is most appropriate for positive chromatography (see Basic Protocol 3 or Alternate Protocol 2).

If positive chromatography is to be performed, continue to use only the immobilized dye column selected for positive chromatography. This column should still retain the desired protein.

Elute desired protein from positive chromatography column

10. Wash the immobilized dye column whose flowthrough contained the least amount of desired protein (i.e., the column selected for positive chromatography) with 2 bed volumes of application solvent containing the lowest concentration of the elution reagent selected for testing. Collect the flowthrough in a separate container.

    A typical nonspecific elution reagent is NaCl, beginning with a lowest concentration of 0.1 M. A typical specific elution reagent is a substrate, cofactor, inhibitor, effector, or ligand, beginning with a lowest concentration of 1 μM. The specific elution reagents chosen for testing should represent a good compromise between availability, economy, and high affinity for the desired protein.

11. Measure the total protein and the total desired protein in the flowthrough.

12. Increase the concentration of the elution reagent in the application solvent by increments and repeat steps 10 and 11.

    A representative increment for a nonspecific elution reagent is 0.2 M and that for a specific elution reagent is an order of magnitude.

    The application solvent containing the lowest concentration of a reagent that elutes at least 80% of the desired protein retained by the immobilized dye column is defined as the elution solvent.

    If alternative elution reagents are to be investigated for reasons of economy or availability, proceed with steps 13 to 15.

Repeat elution using alternative elution solvent (optional)

13. Wash the immobilized dye column selected for positive chromatography with 2 bed volumes of 2 M NaCl followed by 2 bed volumes of application solvent and discard the flowthrough.

14. Reapply a volume of the protein mixture as described in step 6 and wash with two column volumes of application solvent. Discard the flowthrough.

15. Repeat steps 10 to 12 using alternative elution reagent.

ALTERNATE PROTOCOL 1

SELECTION OF COMPONENTS USING CENTRIFUGATION

The immobilized dye for negative or for positive chromatography and the eluant for positive chromatography can also be selected using a centrifugation procedure instead of the chromatography procedure described in Basic Protocol 1. A different immobilized dye is placed in a series of microcentrifuge tubes. The supernatant solution is discarded from each tube and replaced with a small volume of the protein mixture. The contents of each tube are mixed and centrifuged and the supernatants analyzed for the amount of total protein and desired protein. The immobilized dye whose supernatant contains the least amount of total protein and the greatest amount of desired protein is most appropriate for negative chromatography. Conversely, the immobilized dye whose supernatant contains the greatest amount of total protein and the least amount of desired protein is appropriate for positive chromatography. To select an elution reagent for positive chromatography, the same amount of the immobilized dye selected for positive chromatography is added to a series of microcentrifuge tubes. The protein mixture is added to each tube and the

BCU IASH CENTRAL UNIVERSITY LIBRARY

supernatant discarded. A different concentration of a potential elution reagent is added to each tube, mixed, and centrifuged, and each supernatant analyzed for the amount of total protein and desired protein present. Any number of different potential specific and nonspecific elution reagents can be evaluated in this manner. The concentration of elution reagent that provides the greatest-fold purification, recovery of desired protein, and economy of use is selected. The centrifugation procedure allows an investigator to select an elution solvent more expeditiously than does the chromatography procedure because many samples can be processed simultaneously. However, a larger amount of each immobilized dye is required. For a list of materials, also see Basic Protocol 1.

Prepare the protein sample and set up immobilized dye series

1. Prepare protein sample for application to column (Basic Protocol 1, steps 1 to 3).
2. Add a sufficient volume of an immobilized dye suspension to a capped graduated 1.5-ml disposable plastic microcentrifuge tube to bring the solid/liquid boundary to the 0.4-ml mark. Repeat for each available suspension of immobilized dye. Cap each tube and centrifuge at full speed for 1 to 2 min and discard the supernatant.

   If one or more of the immobilized dyes are available in powdered form, suspend about 0.5 g of each powdered dye in application solvent for at least 30 min with gentle stirring prior to adding to the microcentrifuge tubes.

3. Add 1 ml application solvent to each microcentrifuge tube. Cap each microcentrifuge tube and mix its contents by inversion, shaking, or vortexing. Let stand for 30 min to swell the immobilized dye matrix.
4. Centrifuge each tube 1 to 2 min at full speed and discard the supernatant.
5. Add 1 ml application solvent to each tube. Cap each microcentrifuge tube and mix its contents by inversion, shaking, or vortexing. Centrifuge each tube 1 to 2 min at full speed and discard the supernatant.

Apply protein sample to immobilized dye tubes

6. Add the same volume of the protein mixture to each tube.

   The volume of protein mixture should contain between 2.5 and 20 mg of total protein and sufficient desired protein for convenient measurement.

7. Cap each microcentrifuge tube and mix its contents by inversion, shaking, or vortexing. Centrifuge each tube 1 to 2 min at full speed.
8. Measure the amount of total protein and desired protein in each supernatant.

   The immobilized dye whose supernatant contains the least amount of total protein and the greatest amount of desired protein is most appropriate for negative chromatography (see Basic Protocol 2). The immobilized dye whose supernatant contains the greatest amount of total protein and the least amount of desired protein is most appropriate for positive chromatography (see Basic Protocol 3 or Alternate Protocol 2).

Test application solvents for positive chromatography

9. If positive chromatography is to be performed, add a sufficient volume of the immobilized dye suspension selected for positive chromatography to a series of graduated microcentrifuge tubes to bring the solid/liquid boundary in each tube to the 0.4-ml mark.
10. Repeat steps 3 to 6.
11. Construct a series of application solvents each containing concentration of the elution reagent to be tested.

BOU-ASIT-ENTRA- UNIVERSITY LIBRARY

A typical nonspecific elution reagent is NaCl, beginning with a lowest concentration of 0.1 M and increasing in 0.2 M increments. A typical specific elution reagent is a substrate, cofactor, inhibitor, effector, or ligand having a lowest concentration of 1 μM and increasing in order of magnitude increments.

12. Add 1 ml of each application solvent containing the elution reagent to a different centrifuge tube containing immobilized dye. Cap the microcentrifuge tube and mix the contents by inversion, shaking, or vortexing. Centrifuge each tube 1 to 2 min at full speed.

13. Measure the concentration of total protein and of desired protein in each supernatant.

    The application solvent containing the lowest concentration of a reagent that elutes at least 80% of the desired protein retained by the immobilized dye column and that represents a good choice with regard to cost and availability is defined as the elution solvent.

14. Repeat steps 11 to 13 to test additional elution reagents or to refine the concentration of the selected elution reagent, if desired.

BASIC PROTOCOL 2

NEGATIVE CHROMATOGRAPHY

A protein mixture is applied to an immobilized dye column selected for negative chromatography and the column is washed with application solvent. Many of the undesired proteins are retained by the column while the desired protein as well as some of the undesired proteins flow through the column. The concentrations of the total protein and total desired protein in the combined flowthrough are measured. These values should reflect a yield of >80% of desired protein and a modest-fold purification. Negative chromatography is the simplest to execute but, because the desired protein as well as some undesired proteins are not retained and appear in the column flowthrough, it is the least likely to dramatically increase the fold purification of a desired protein. However, it is particularly convenient for the rapid removal of troublesome contaminating proteins such as proteases, nucleases, or serum albumin.

Materials

Protein mixture to be purified
Application solvent, with and without 2 M NaCl
Immobilized dye selected for negative chromatography (see Basic Protocol 1 or Alternate Protocol 1)
0.02% (w/v) sodium azide

Empty chromatographic column with height/diameter ratio ≤5 (sintered glass funnel can suffice)

1. Prepare the protein mixture for chromatography (see Basic Protocol 1, steps 1 to 3).

2. Divide the total protein concentration (in milligrams) to be purified by 5 to estimate the bed volume (in milliliters) of the immobilized dye needed. Add suspension of the immobilized dye to the chromatographic column until that bed volume is acquired.

   Division is by 5 because it is conservatively estimated that each milliliter of column bed volume can retain 5 mg of protein.

   Immobilized dyes may be purchased either as suspensions or as dry powders from commercial sources (see Table 9.2.1). The suspensions are easier to use because they are already swollen; therefore they are the form of choice. If the powdered form of the immobilized dye is used, it should be suspended in an application solvent for about an hour with gentle stirring prior to pouring the column.

BCU IASI CENTRAL UNIVERSITY LIBRARY

3. Wash the column with application solvent until no further dye is eluted.

   The volume of application solution required for washing depends upon the storage of the column. The column matrix slowly but steadily hydrolyzes, generating soluble matrix fragments still containing covalently attached dye molecules. The longer the storage time and the higher the temperature, the greater the amount of solubilized dye. The volume of wash is determined by the investigator; when colored dye stops appearing in the wash, the column is ready for chromatography. This may require 1 or 20 column volumes.

   The presence of the colored dye can be detected in the washing either by visual inspection or by spectrophotometric measurement at an appropriate wavelength.

4. Apply the entire protein mixture to be purified to the immobilized dye column using a flow rate of at least 1 ml/min. Collect the flowthrough liquid in a single container.

   The flow rate is dictated by the size of the column and the density of its packing. A fast flow rate reduces experimental time. Equilibration of protein with immobilized dye is rapid; there is little advantage in using flow rates less than 1 ml/min.

5. Wash the column with 1 column volume of application solvent. Collect the flow-through liquid in the same container as in step 4.

6. Measure the concentration of total protein and total desired protein in the combined flowthrough liquids.

   These values should reflect a good yield, >80% of desired protein, and a modest-fold purification. If the fold purification is less than expected from the results observed using Basic Protocol 1 or Alternate Protocol 1, add more immobilized dye to the column and reapply the combined flowthrough liquids.

7. Wash the column with 1 column volume of application solvent containing 2 M NaCl to displace the retained protein.

   Protein displacement can be ascertained by monitoring the amount of protein in the wash effluent.

8. Prepare the column for storage by washing with 2 column volumes of 0.02% sodium azide and storing the poured column at 4°C.

POSITIVE CHROMATOGRAPHY USING BATCH ELUTION

BASIC PROTOCOL 3

A protein mixture is passed through an immobilized dye column selected for positive chromatography. This immobilized dye retains the desired protein and some of the undesired proteins. The desired protein is then selectively eluted using either a nonspecific or a specific elution reagent (a higher-fold purification will likely be obtained using the specific eluant), and the concentrations of total protein and desired protein in each fraction are measured. This protocol for positive chromatography utilizes batch elution, which is simpler to execute than Alternate Protocol 2 utilizing gradient elution (the latter may yield a higher-fold purification than batch elution but requires more effort to set up).

Materials

Protein mixture to be purified
Appropriate application solvent
Immobilized dye and elution solvent selected for positive chromatography (see Basic Protocol 1 or Alternate Protocol 1)
Empty chromatographic column with height/diameter ratio >5
Fraction collector

1. Prepare protein sample and immobilized dye column (see Basic Protocol 2, steps 1 to 5) using the immobilized dye selected for positive chromatography.

BOSTON COLLEGE CENTRAL UNIVERSITY LIBRARY

2. Measure the concentration of the total protein and the desired protein in the flow-through liquids. If the majority of the desired protein is retained by the column, discard the flowthrough liquid. If a significant amount of desired protein is not retained, add more immobilized dye to the column and reapply the flowthrough liquid.

   Typically, at least 70% and, if possible, 90% or more of the desired protein is retained.

3. Wash the column with the elution solvent selected for positive chromatography using a flow rate of at least 1 ml/min. Collect the flowthrough in a fraction collector in constant-volume fractions, typically 1 ml.

4. Measure the concentration of total protein and desired protein in each fraction.

5. Pool fractions containing the desired protein.

   The specific activity of the desired protein in the fractions can be used to determine which fractions to include in the pool. Typically pool all tubes containing at least 10% of the total amount of desired protein applied to the column.

6. Wash and store the column (see Basic Protocol 2, steps 7 and 8).

ALTERNATE PROTOCOL 2

POSITIVE CHROMATOGRAPHY USING GRADIENT ELUTION

A protein mixture is passed through an immobilized dye column selected for positive chromatography. This immobilized dye retains the desired protein and some of the undesired proteins. The desired protein is selectively eluted using a linear gradient constructed from the elution solvent selected for positive chromatography, and the concentrations of total protein and desired protein in each fraction are measured. This protocol is more elaborate to set up and operate, but sometimes generates a greater-fold purification than obtained using the batch elution system of Basic Protocol 3.

***Additional Materials** (also see Basic Protocol 3)*

Empty chromatographic column having a height/diameter ratio of >5
A coated magnet and magnetic stirrer
Linear gradient maker (commercial or constructed from two beakers of the same size connected by a siphon)

1. Prepare protein sample and immobilized dye column (see Basic Protocol 3, steps 1 and 2).

   All solutions should be carefully introduced so that the column packing is minimally disturbed.

2. Place a volume of the elution solvent selected for positive chromatography in the nonmixing reservoir of the linear gradient maker. This volume is typically equal to the column volume.

   The nonmixing reservoir is the reservoir more distant from the column. Its contents flow into the mixing reservoir prior to introduction onto the column.

3. Place 10% of the volume of elution solvent used in step 2 in the mixing reservoir of the gradient maker. Add application solvent to make up the remaining 90% of the volume.

   This procedure will generate an order-of-magnitude gradient terminating in the eluant concentration employed in batch elution.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

4. Place a magnetic stir bar in the mixing reservoir. Place the mixing reservoir on a magnetic stirrer. Position the nonmixing reservoir so that the bottom of this vessel is adjacent to and coplanar with the bottom of the mixing reservoir.
5. Provide for liquid flow between the two reservoirs and between the mixing reservoir and the column.
6. Gently mix the liquid in the mixing reservoir. Initiate liquid flow between the two reservoirs and between the mixing reservoir and the column. Use a flow rate of 1 ml/min and collect 1-ml fractions.

   Resolution of desired and undesired proteins may be improved by narrowing the concentration range of the specific or nonspecific elution reagent, by increasing the total volume of the elution solutions, or by decreasing the flow rate of the elution solutions.
7. Measure the protein concentrations, pool the appropriate fractions, and wash the column (see Basic Protocol 3, steps 4 to 6).

B. C. U. „M. Eminescu" IAŞI
CHIMIE

TANDEM CHROMATOGRAPHY

BASIC PROTOCOL 4

In tandem chromatography, columns for negative and positive immobilized dye chromatography are sequentially linked. The protein mixture is first passed through the column containing the immobilized dye chosen for negative chromatography, which preferentially removes undesired protein. The effluent from this column is immediately passed over a second column containing the immobilized dye chosen for positive chromatography, which retains the desired protein. The desired protein is then selectively eluted from the second column using either batch or gradient elution and the elution solvent selected for positive chromatography.

Materials

Protein mixture to be purified
Appropriate application solvent
Immobilized dyes selected for negative and for positive chromatography (Basic Protocol 1 or Alternate Protocol 1)
Selected elution solvent
2 M NaCl
0.02% (w/v) sodium azide

Empty chromatographic column with height/diameter ratio >5 (two columns are required if positive chromatography is to be performed using batch elution)
Empty chromatographic column with height/diameter ratio ≥5 (if positive chromatography is to be performed using gradient elution)
Linear gradient maker (for gradient elution)
Fraction collector
Pipettors and disposable tips

Prepare the protein sample and columns

1. Prepare the protein mixture for chromatography (see Basic Protocol 1, steps 1 to 3).
2. Prepare the negative and positive columns (see Basic Protocol 2, steps 2 and 3).

   The immobilized dye selected for negative chromatography is placed in a column with a height/diameter ratio ≤5. The immobilized dye selected for positive chromatography is placed in another identical column (if batch elution is to be employed) or in a column with a height/diameter ratio ≥5 (if gradient elution is to be employed).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

3. Position the negative column immediately above the positive column. Connect the two columns so that the effluent of the upper column flows directly on top of the lower column.

Pass sample through columns

4. Apply the protein mixture to the upper negative column. Collect the flowthrough liquid from the lower positive column in a single container.
5. Wash the tandem columns with 1 collective column volume of application solvent. Add the flowthrough from the washing to the container in step 4.
6. Measure the concentration of desired protein in the collective flowthrough to ascertain that the desired protein has been retained.
7. Uncouple the upper negative column and set it aside.

Recover the protein

8. Wash the positive column with 2 column volumes of application solvent.
9. Subject the positive column to either batch or gradient elution using the elution solvent selected for positive chromatography. Perform and process batch elution (see Basic Protocol 3, steps 3 to 5) or gradient elution (see Alternate Protocol 2, steps 2 to 6).
10. Wash and store each column (see Basic Protocol 2, steps 7 and 8).

SUPPORT PROTOCOL

IMMOBILIZATION OF REACTIVE DYES

Reactive dyes must be covalently attached to an insoluble porous matrix ("immobilized") to be used in a chromatographic protocol for protein purification. A reactive dye is a visible chromophore that has a good chemical leaving group, as illustrated in Figure 9.2.1. Immobilization of a reactive dye is a simple procedure that can be performed in the laboratory to generate immobilized dyes that cannot be purchased or to save the cost of commercial immobilization. Table 9.2.2 lists a selection of dyes that can be purchased at modest cost. Each of these dyes can be immobilized by reacting it with a chromatographic matrix in aqueous solvents using moderate conditions. The nature of the covalent attachment is illustrated in Figure 9.2.1B. This protocol is scaled for preparation of an immobilized dye column having a bed volume of 100 ml.

Materials

Reactive dye (Table 9.2.2)
Concentrated solution of KCl
Reactive chromatographic matrix: e.g., Sepharose CL-4B or CL-6B (Pharmacia Biotech) or other cross-linked agarose
4 M and 1 M NaCl
10 M NaOH
2 M NH_4Cl

Large sintered glass funnel
Plastic disposable filter with pore size of 0.45 μm

1. Remove any salt, buffer, or surfactant from the reactive dye sample by precipitation of the dye as the potassium salt from aqueous solution and filtration of the precipitate using a sintered glass funnel mounted on a filter flask. Place several grams of the dye in a beaker and add water to make a concentrated solution of the dye. Add concentrated solution of KCl dropwise to precipitate the dye from the solution. Pass the dye

BCU IASI - CENTRAL UNIVERSITY LIBRARY

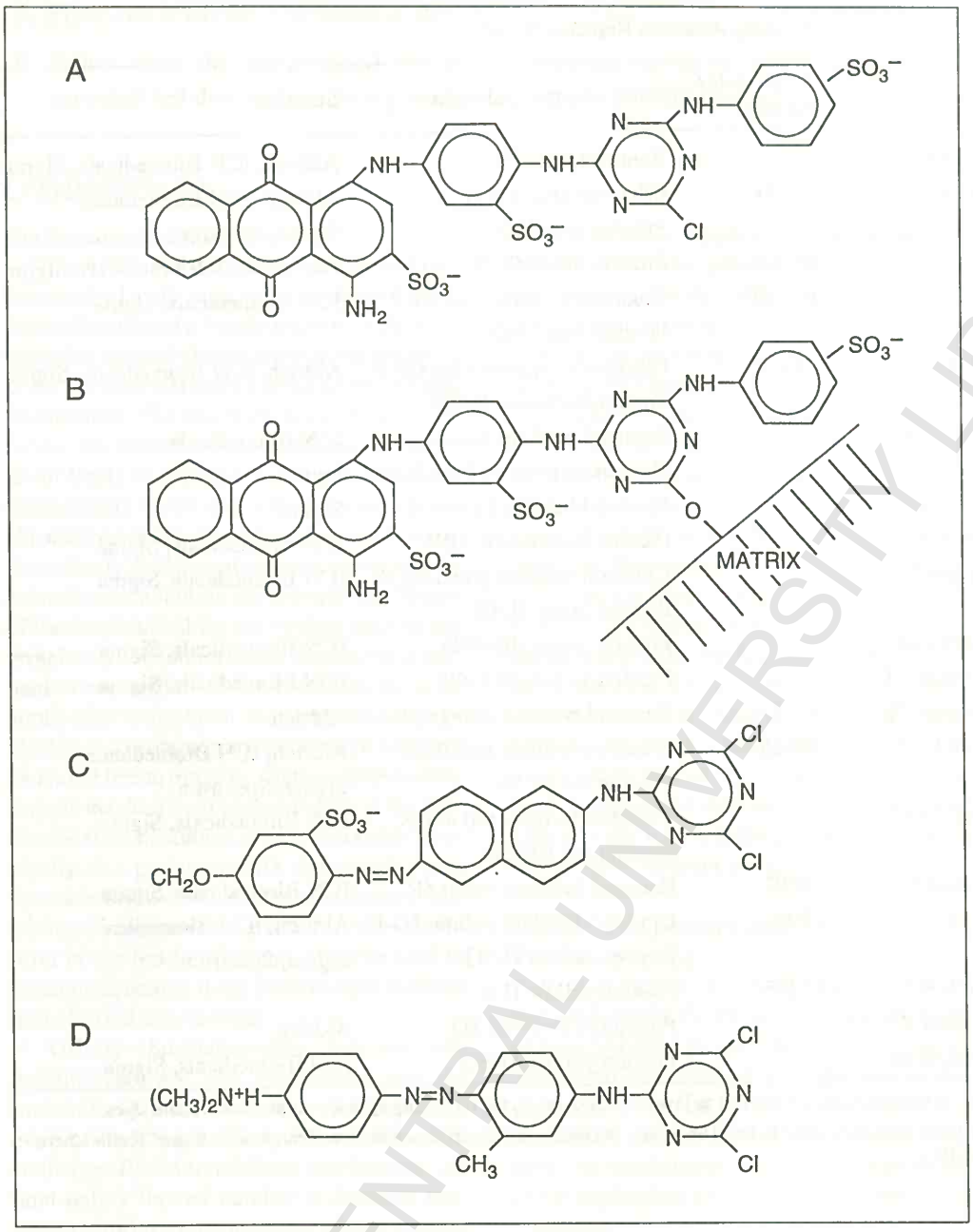


Figure 9.2.1 Structures of some reactive dyes. The dyes are arranged so that their triazine rings, the rightmost rings in structures **C** and **D**, have a common position. (**A**) Reactive blue 2, color index 61211, a reactive monochlorotriazine dye, also designated as Cibacron blue 3G-A and Procion blue H-B. (**B**) Immobilized reactive blue 2, also designated as Affi-Gel blue gel, Blue Sepharose CL-6B, Cibacron blue 3G-A-agarose, and DyeMatrex blue A gel. (**C**) Reactive red 8, color index 17908, a reactive dichlorotriazine dye, also designated as Procion scarlet MX-G. (**D**) A positively charged reactive dye.

suspension through a plastic disposable filter with a pore size of 0.45 µm, wash the filtered precipitate with ~100 ml of water, and air dry the washed precipitate.

The most common reactive dyes—e.g., Procion (ICI) and Cibacron (Ciba-Geigy) dyes—have a sulfonated chromophore linked to a chlorotriazine group by an aminoether bridge.

2. Suspend 80 g chromatographic matrix in 280 ml water.
3. Dissolve 1.2 g reactive dye in 80 ml water and add it to the matrix suspension.

BCIT LIBRARY CENTRAL UNIVERSITY LIBRARY

Table 9.2.2 Some Commercially Available Reactive Dyes[a]

Generic name	Color index number[b]	Some commercial names	Suppliers
Reactive black 5		Remazol black B	Aldrich, ICN Biomedicals, Sigma
Reactive blue 2	61211	Cibacron blue 3G-A	Aldrich, ICN Biomedicals
		Procion blue H-B	Sigma, Spectrum
Reactive blue 4	61205	Procion blue MX-R	Aldrich, ICN Biomedicals, Sigma
Reactive blue 5	61210	Cibacron brilliant blue BR-P	ICN Biomedicals, Sigma
		Procion blue H-GR	
Reactive blue 15	74459	Cibacron turquoise blue GF-P	Aldrich, ICN Biomedicals, Sigma
		Procion turquoise H-GF	
Reactive blue 19		Remazol brilliant blue R	ICN Biomedicals
Reactive blue 114		Drimarene brilliant blue K-BL	Sigma
Reactive blue 160		Procion blue HE-RD	Sigma
Reactive brown 10		Procion brown MX-5BR	ICN Biomedicals, Sigma
Reactive green 5		Cibacron brilliant green 4G-A	ICN Biomedicals, Sigma
		Procion green H-4G	
Reactive green 19		Procion green HE-4BD	ICN Biomedicals, Sigma
Reactive orange 14		Procion yellow MX-4R	ICN Biomedicals, Sigma
Reactive orange 16	17757	Remazol brilliant orange 3R	Aldrich
Reactive red 4	18105	Cibacron brilliant red 3B-A	Aldrich, ICN Biomedicals
		Procion red H-7B	Sigma, Spectrum
Reactive red 120		Cibacron brilliant red 4G-E	ICN Biomedicals, Sigma
		Procion red HE-3B	
Reactive violet 5	18097	Remazol brilliant violet 5R	ICN Biomedicals, Sigma
Reactive yellow 2	18972	Cibacron brilliant yellow 3G-P	Aldrich, ICN Biomedicals
		Procion yellow H-5G	Sigma, Spectrum
Reactive yellow 3	13245	Procion yellow H-A	
Reactive yellow 81		Procion yellow HE-3G	Aldrich
Reactive yellow 86		Procion yellow M-8G	ICN Biomedicals, Sigma

[a]Dichlorotriazine reactive dyes are denoted as Procion MX dyes; all the remaining dyes are monochlorotriazine dyes.

[b]Numbers are taken from the Colour Index, The Society of Dyers and Colourists and the American Association of Textile Chemists and Colorists (1971).

4. Add 40 ml of 4 M NaCl to the matrix suspension.
5. If a monochlorotriazine reactive dye is used, add 4 ml of 10 M NaOH to the matrix suspension and gently stir it with a magnetic stirrer and stir bar for 72 hr at ambient temperature or 16 hr at 55° to 60°C. If a dichlorotriazine reactive dye is used, add 0.5 ml of 10 M NaOH and gently stir the matrix suspension with a magnetic stirrer and stir bar for 4 hr at ambient temperature.
6. Filter the suspension using the sintered glass funnel and wash the solid material on the filter with copious quantities of water, then 1 M NaCl, and then water again until the filtrate is clear.
7. Suspend the material on the filter in 2 M NH_4Cl, pH 8.5, and gently stir with a magnetic stirrer and stir bar for 4 hr at ambient temperature.

   This step eliminates any remaining chloro groups.

BCU IASI CENTRAL UNIVERSITY LIBRARY

8. Filter and wash the suspension as described in step 6.
9. Either store the immobilized dye as a suspension in water at 4°C or dry the immobilized dye on a sintered glass filter and store as a dry powder.

COMMENTARY

Background Information

The surfaces of all proteins bind at least one biochemical with high affinity. Accordingly, immobilization of a biochemical on a chromatographic support should result in the retention of all proteins that have a binding site for that biochemical. The biochemical can exist in two forms: the mobile biochemical, which moves about freely in solution, and the immobilized biochemical, which is covalently attached to a chromatographic support and cannot move about freely. In chromatography, a competition exists between mobile and immobilized forms of the biochemical for the binding sites on the proteins. If the immobilized form is the dominant concentration of the biochemical, as in application to a column, the protein exists principally as a protein/immobilized dye complex and is not free to move in solution. If the mobile form is the dominant concentration of the biochemical, as in elution, the protein exists principally as a protein/mobile dye complex and moves freely in solution. Elution solutions containing a concentration gradient of the mobile form of the biochemical can often be used to distinguish among those proteins that bind the immobilized biochemical.

Affinity chromatography, however, can present problems. The biochemical must be immobilized in such a manner that the protein can bind it with high affinity, often a substantial challenge. Each immobilized biochemical can bind only a limited number of different proteins, requiring construction of a large number of different affinity columns for purification of a series of proteins. Further, the degradative enzymes present in crude cellular extracts often cleave immobilized biochemicals into useless remnants, necessitating either the frequent replacement of affinity columns or the postponement of their use until the end of a purification scheme. Such postponement compromises the selective power of affinity chromatography, which should, in principle, purify a protein to homogeneity in a single step.

Fortunately, immobilized dye affinity chromatography circumvents these problems. A single reactive dye can effectively bind a large number of different biochemicals. A reactive dye is sufficiently complex that immobilization does not reduce its affinity. An immobilized reactive dye is not degraded by enzymes commonly found in crude protein mixtures, as the chemistry of the dye is quite distinct from that of typical biochemicals.

An immobilized dye functions as a stable general affinity molecule that can be made selective by the nature and concentration of the mobile biochemical chosen as the elution reagent. A single immobilized dye, variously known as Cibacron blue (F)3G-A, Procion blue H-B, reactive blue 2, blue dextran, Affi-Gel blue, or DyeMatrex blue A, has contributed significantly to the purification of over 60 different proteins. Some of these proteins have been purified to homogeneity from a crude bacterial extract in a single step providing a several-thousand-fold purification with >80% yield. Immobilized dyes as a group have contributed significantly to the purification of several hundred different proteins.

Critical Parameters and Troubleshooting

Several aspects of dye affinity chromatography are critical. Foremost is the ionic strength of the application solvent. Because application solvents of ionic strength >50 mM weaken the electrostatic interactions that contribute to the binding of desired proteins by immobilized dyes, the concentrations of ionic components in the application solvent should be kept to a minimum.

Second is the presence of nonionic detergents in the application solvent. Nonionic detergents are frequently used to solubilize intrinsic protein located in subcellular organelles such as membranes. These nonionic detergents form micelles that encapsulate immobilized dyes, making them inaccessible to proteins. Addition of a modest concentration of a negatively charged detergent, such as SDS or deoxycholate, to an application solvent containing a nonionic detergent results in the formation of mixed micelles. The negatively charged immobilized dyes repel the negatively charged mixed micelles in solvents of modest ionic strength, preserving the accessibility of

BAYLOR COLLEGE OF MEDICINE LIBRARY

immobilized dyes to desired proteins (Robinson et al., 1980). The nonionic and ionic detergents should have a concentration ratio of ~5:1.

Third is the sign of the electrostatic charges on immobilized dyes. Most reactive dyes are negatively charged; this is normally an advantage, as most protein binding sites contain positively charged amino acid residues designed to bind negatively charged biochemicals, such as cAMP, glucose 6-phosphate, or oligo(dAT). However, some biochemicals have a positive charge; the proteins to which they bind must then contain negatively charged residues. These binding proteins are electrostatically repelled by negatively charged immobilized dyes in solvents of modest ionic strength. The immobilized form of a positively charged dye (Fig. 9.2.1D) retains proteins having binding sites for positively charged biochemicals, such as trypsin, thrombin, and carboxypeptidase (Clonis et al., 1987). This immobilized dye is particularly useful in removing contaminating proteases from cellular extracts by negative chromatography.

The two primary problems encountered in immobilized dye affinity chromatography are the inability to retain a desired protein and the inability to elute a desired protein that has been retained. The inability to retain a desired protein may result from a number of situations. If a fresh immobilized dye column is being used, the following may pertain:

The ionic strength of the protein mixture may be too high. Either dilute the protein mixture with distilled water or dialyze it against a low ionic strength application solvent, such as 50 mM Tris·Cl devoid of neutral salts.

The concentration in the protein mixture of the biochemical that binds to the desired protein may be too high. Either dilute the mixture with distilled water or, if the biochemical is sufficiently small, dialyze the mixture against a low ionic strength application solvent.

Retention of the desired protein may necessitate the presence of a metallic cation, such as Mg^{2+} or Zn^{2+}. Add a minimal concentration of one or more likely metallic cations to the application solvent.

A nonionic detergent may be present in the protein mixture. Add a sufficient amount of a negatively charged detergent, such as SDS or deoxycholate, to the crude protein mixture to maintain the accessibility of the immobilized dye. However, note that SDS, even in low concentrations, denatures many proteins.

The immobilized dye may have a relatively weak affinity for the desired protein. Add more immobilized dye to the column or screen a series of alternative immobilized dyes for one having a stronger affinity for the desired protein.

The desired protein is designed to bind only positively charged biochemicals. Purchase or synthesize an immobilized dye having a net positive charge.

The desired protein may be among a small group of proteins that is not retained by any immobilized dye. Negative chromatography can then be used to great advantage, particularly a series of negative columns in tandem.

If an old immobilized dye column is being used, the following may pertain:

Insufficient immobilized dye may be accessible. After repeated use of an immobilized dye column the dye may become bound with proteins, detergents, and other components in cellular extracts that are not removed by washing with 2 M NaCl. An immobilized dye column should be periodically washed with several column volumes of one or more of the following solutions to displace these dye-bound components: a strong denaturant such as 8 M urea, 6 M guanidine·HCl, or 1% SDS; a potent lyotropic reagent such as 3 M KSCN; a strong base such as 0.5 M NaOH; or a mixed solvent such as chloroform/methanol.

Insufficient immobilized dye may be present. If an immobilized dye column has been used extensively or stored in solution, the glycosidic bonds linking the sugars in the agarose matrix will be hydrolyzed. Such hydrolysis results in loss of the immobilized dye. A new immobilized dye column is strongly colored. A depleted immobilized dye column is lightly colored and should be replaced.

The inability to elute a desired protein that has been retained on an immobilized dye column presents a different problem, in that the affinity is too strong rather than too weak. Suggestions for weakening the affinity include the following:

Performing the chromatography at a different pH. Most proteins exhibit a pH optimum for binding a biochemical. This pH optimum results in part from changes in the ionic character of some of the amino acid side chains in the protein that are directly involved in binding. Changes in pH that weaken the affinity of biochemicals are likely to weaken the affinity for the immobilized dye as well. Use Basic Protocol 1 to determine the effect of pH on the elution of the desired protein. Note, however, that pH values outside the range 4 to 11 may result in irreversible inactivation.

PCU IASTRI PATNA UNIVERSITY LIBRARY

Performing the chromatography at a different temperature. The binding of a biochemical by a protein involves a combination of weak noncovalent bonds including hydrogen bonding, hydrophobic interactions, and electrostatic interactions. Changes in temperature weaken some of these bonds and strengthen others. Accordingly, changing the temperature may weaken the affinity of the immobilized dye for the desired protein. Note, however, that the temperature should not be increased above 50°C to prevent thermal inactivation of the desired protein.

Increasing the potency of the elution solvent. Nonspecific elution solvents promote elution by weakening the electrostatic and hydrophobic contributions to affinity. The electrostatic contributions will usually be minimized in a 0.2 M solution of any neutral salt. However, minimization of the hydrophobic contributions requires higher concentrations and judicious choice of the neutral salt. The lyotropic or Hofmeister series reflects the relative potency of neutral salts. A 3 M solution of the neutral salt KSCN is among the most potent for weakening hydrophobic interactions (Robinson et al., 1981).

Specific elution solvents, by contrast, weaken affinity by competing with the immobilized dye for the desired protein. The higher the concentration of the biochemical and the stronger its affinity for the desired protein, the greater its effectiveness as an eluant. Accordingly, the concentration of a specific elution solvent should be increased in an effort to elute the desired protein. Clearly, a limiting concentration will be set by the solubility, availability, and cost of the specific eluant. Alternatively, other biochemicals known to bind to the desired protein as substrates, coenzymes, cofactors, competitive inhibitors, and allosteric effectors may be surveyed for their efficacy as elution reagents.

Adding a chelation agent. Metallic cations often function as obligatory cofactors in the binding of biochemicals to proteins. Accordingly, the affinity of the desired protein for the immobilized dye may be weakened by the presence of a chelation agent such as 0.1 M EDTA or EGTA.

Combining specific elution reagents. Immobilized dyes are sufficiently large that a single dye molecule may occupy adjacent biochemical binding sites on a protein: e.g., a dye may occupy the substrate site and the coenzyme site on an enzyme surface. Alternatively, two different immobilized dye molecules may simultaneously bind to two remote sites on the same protein molecule: e.g., different immobilized dye molecules may simultaneously bind to the substrate site and to the allosteric site on an enzyme. In each of these situations, the elution of a desired protein can be accomplished using a combination of specific elution reagents.

Adding a protein denaturant such as 6 M guanidine·HCl. Most proteins can be eluted from an immobilized dye column using a denaturant that unfolds protein binding sites. Because most retained proteins have already been eluted, addition of a denaturant should result in a substantial-fold purification, provided the desired protein is denatured reversibly. Such reversibility is commonly achieved by dilution or dialysis of the column eluate with application solvent.

Lowering the density of the immobilized dye. Many desired proteins are oligomers containing two or more identical binding sites. If the chromatographic matrix contains a high density of accessible immobilized dye molecules, it is likely that two or more immobilized dye molecules will simultaneously bind to a single desired oligomeric protein (Hogg and Winzor, 1985). Such multiple binding makes elution more difficult. Diluting an immobilized dye matrix with an unmodified matrix will not decrease the density of the immobilized dyes on each chromatographic bead. The only remedy is to obtain the same immobilized dye from an alternative supplier, hoping to obtain a less dense sample, or to immobilize the dye in the laboratory using the Support Protocol.

Selecting an alternative immobilized dye that has a weaker affinity for the desired protein. Such weakening should facilitate improved results in positive chromatography.

Anticipated Results

The yield of the desired protein following immobilized dye chromatography is commonly >80%. Random differences in yield occur depending upon how completely the desired protein is retained (positive chromatography) or not retained (negative chromatography) and how completely the retained protein is eluted. Lesser yields are usually the result of proteolysis either before or during chromatography. Proteolysis may generate remnants of the desired protein that have weaker affinities for an immobilized dye. Such remnants elute prior to the elution of the dominant unproteolyzed desired protein and would not be included in the pooled fractions. Good yields of

EXPASY CENTRAL WEBSITE LIBRARY

desired proteins result in part from the speed of affinity chromatography, the temperature, the stabilization provided to the desired protein from binding an immobilized dye, and the respectful treatment of a crude protein mixture by an experienced investigator.

The fold purification achieved by immobilized dye chromatography is largely controlled by the relative population of the desired protein in a protein mixture. If 10% of the mixture is the desired protein, then a 10-fold purification step is maximal. If <0.1% of the mixture is the desired protein, then a >1000-fold purification step is feasible and has been observed. However, in most instances 10- to 100-fold purifications are observed.

Time Considerations

The preliminary survey can be completed in a few hours or a day, depending on its breadth. Once a protocol is selected, negative chromatography should be completed in 1 hr and positive chromatography in a few hours. These time estimates greatly depend on the volume of the solutions to be processed and the flow rates employed.

Literature Cited

Clonis, Y.D., Stead, C.V., and Lowe, C.R. 1987. Novel cationic triazine dyes in protein purification. *Biotechnol. Bioeng.* 30:621-627.

Hogg, P.J. and Winzor, D.J. 1985. Effects of solute multivalency in quantitative affinity chromatography: Evidence for cooperative binding of horse liver alcohol dehydrogenase to blue Sepharose. *Arch. Biochem. Biophys.* 240:70-76.

Robinson, J.B., Jr., Strottmann, J.M., Wick, D.G., and Stellwagen, E. 1980. Affinity chromatography in nonionic detergent solutions. *Proc. Natl. Acad. Sci. U.S.A.* 77:5847-5851.

Robinson, J.B., Jr., Strottmann, J.M., and Stellwagen, E. 1981. Prediction of neutral salt elution profiles for affinity chromatography, *Proc. Natl. Acad. Sci. U.S.A.* 78:2287-2291.

The Society of Dyers and Colourists and the American Association of Textile Chemists and Colorists. 1971. Colour Index, 3rd edition. The Society of Dyers and Colourists, Bradford, England, and the American Association of Textile Chemists and Colorists, Research Triangle Park, NC.

Key References

Clonis, Y.D., Atkinson, T., Bruton, C.J., and Lowe, C.R. 1987. Reactive Dyes in Protein and Enzyme Technology. Macmillan, New York.

A comprehensive review of the preparation, immobilization, and utilization of reactive dyes for protein purification.

Lowe, C.R., Burton, S.J., Burton, N., Stewart, D.J., Purvis, D.R., Pitfield, I., and Eapen, S. 1990. New developments in affinity chromatography. *J. Mol. Recognit.* 3:117-122.

A lively discussion of emerging developments.

Scopes, R.K. 1994. Protein Purification, 3rd ed. Springer-Verlag, New York.

A useful discussion of immobilized dye affinity chromatography by the advocate of tandem chromatography.

Stellwagen, E. 1990. Chromatography on immobilized reactive dyes. *Methods Enzymol.* 182:343-357.

An alternative discussion of the topic with more emphasis on strategy and less on detail.

Stellwagen, E. 1993. Affinity chromatography with immobilized dyes. *In* Molecular Interactions in Bioseparations (T.T. Ngo, ed.) pp. 247-255. Plenum Press, New York.

A detailed review of the interaction of reactive blue 2 with a typical oligomeric protein, alcohol dehydrogenase.

Contributed by Earle Stellwagen
University of Iowa
Iowa City, Iowa

BCU IASI / CAMPUS UNIVERSITY LIBRARY

Affinity Purification of Natural Ligands

Immobilization of proteins, nucleic acids, and other "bioligands" is not simple. The proper bioligand needs to be selected for the specific application of the product. This decision influences all others: the matrix is chosen and activated by the method that is most appropriate for this specific application and the ligand is coupled under conditions dictated by the activation method and the nature of the ligand (e.g., is it a labile protein or a sturdy enzyme cofactor?). There are many matrices and activation and coupling methods, and new ones are constantly being developed. The most frequently used are hydrophilic matrices, which exhibit the necessary low nonspecific binding, and coupling systems that react primarily with nucleophilic residues, such as lysyl amine groups in proteins. Orientation of the resulting biomolecule is extremely important, although these effects have been only superficially investigated. Immobilized IgG is only twofold more efficient when immobilized solely by the Fc portion (Murayama et al., 1978; Doman et al., 1990; Hermanson et al., 1992), whereas the interactions of large proteins that bind macromolecular ligands (lysozyme is a good example) are changed by several orders of magnitude depending on the orientation of the molecule (Fig. 9.3.1). For this reason, it is fortunate that both cationic (e.g., fluoromethyl pyridinium) and hydrophobic (e.g., Emphaze, tosyl agarose) activation methods are available.

The choice of matrix and the stability of the bound ligand determine the choice of activation and coupling methods. Preactivated matrices are available but are very costly. Therefore they are best used when small amounts of ligand are to be immobilized for "test" columns and methods development. Larger projects and large-scale applications require ligand activation in the laboratory. Because of their frequent use in the past, agarose gel is often the matrix of choice and cyanogen bromide is often the activation and coupling agent of choice. However, depending on the procedure, neither may be optimal and before beginning a project it is best to explore the range of matrices and methods carefully, particularly if later industrial level scale-up with regulatory implications is anticipated. An overview of the major activation chemistries is given in Tables 9.3.1 and 9.3.2. Of the many activation systems available, three of the most frequently used are

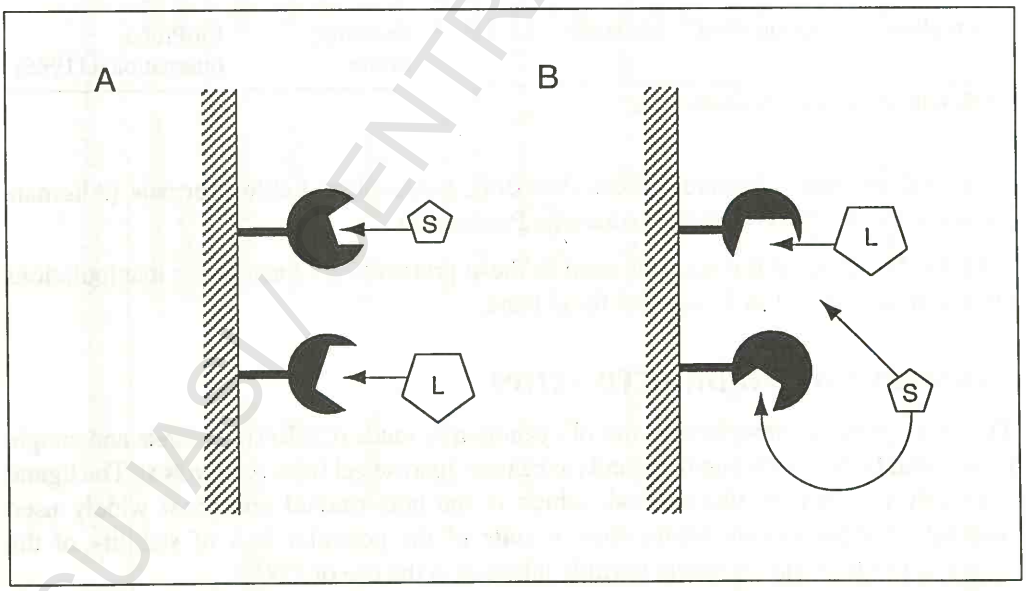


Figure 9.3.1 Binding of small (S) versus large (L) substrates to an oriented (A) and randomly (B) immobilized enzyme. Arrows signify the movement of the enzyme to the ligand and its subsequent binding to the ligand. Reproduced from Voivodov et al. (1993) with permission of Hüthig & Wepf Publishers, Zug, Switzerland.

Copyright © 1995 by John Wiley & Sons, Inc.

FCU IAS / CENTRAL UNIVERSITY LIBRARY

Table 9.3.1 Selected Reagents to Activate Matrix Hydroxyl Functions[a]

Activation method	Stability of bond	Group that reacts	Reagent toxicity	Activation time (hr)	Type of bond	References
Tresyl chloride, sulfonyl chlorides	Excellent	Thiols, amines	Low	0.1-1.0	Secondary amine	Lawson et al. (1983); Gribnau (1977)
Cyanogen bromide	Poor	Amines	High	0.1-0.4	Isourea or imidocarbonate	Axen et al. (1967)
Bisoxiranes (epoxides)	Excellent	Thiols	Moderate	5-18	Secondary amine	Axen et al. (1967)
Epichlorohydrin	Excellent	Thiols, amines	Moderate	2-24	Secondary amine	Axen et al. (1967)
Triazine	Good	Amines, thiols, hydroxyls	High	0.5-2	Triazine ether	Gribnau (1977); Lily (1976)
Benzoquinone	Good	Amines	Low	1-2	Anilinyl	Axen et al. (1967)
Divinyl sulfone	Poor in base, fair at pH <7.0	Amines	High	0.5-2	Secondary amine	Axen et al. (1967)
Glutaraldehyde	Excellent	Amine	Moderate	5-18	Secondary amine (?)	Axen et al. (1967)
Phosphoryl chloride	Excellent	Amine	High	18	Amide	Ngo and Lenhoff (1980)
p-Nitrophenyl chloroformate	Good	Amine	Moderate	1.0	Urethane	Wilchek and Miron (1982); Drobnik et al. (1982)
N-Hydroxysuccinimidyl chloroformate	Good	Amine	Moderate	1.0	Urethane	Wilchek and Miron (1982); Drobnik et al. (1982)
Carbonyldiimidazole	Good	Amine	Moderate		Urethane	Bethell et al. (1979)
1-Cyano-4-dimethyl-aminopyridinium tetrafluoroborate	Poor	Amines	Moderate	20 min	Isourea	Kohn and Wilchek (1984)
2-Fluoro-1-methyl-pyridinium tosylate	Excellent	Amine/thiol	Moderate	0.5-1.0	Secondary amine	BioProbe International (1986)

[a]Reproduced from Scouten (1987) with permission of Academic Press.

described: cyanogen bromide (Basic Protocol), *p*-nitrophenyl chloroformate (Alternate Protocol 1), and tresyl chloride (Alternate Protocol 2).

CAUTION: Many of the reagents used in these protocols are hazardous; manipulations should be performed in a chemical fume hood.

BASIC PROTOCOL

CYANOGEN BROMIDE ACTIVATION

This basic protocol describes the use of cyanogen bromide (CNBr) to activate and couple proteins and other nucleophilic ligands to beaded agarose gel (also see *UNIT 9.5*). The ligand is readily coupled by this method, which is the best studied and most widely used technique for protein immobilization in spite of the potential lack of stability of the resulting product and the safety hazards inherent in the use of CNBr.

CAUTION: CNBr is very toxic, releases toxic vapor, and is reported to create explosive side products upon storage; therefore, it should always be used in a hood and handled with appropriate CNBr- and solvent-resistant gloves. Only a fresh bottle of CNBr should be

BCU IASI CENTRAL UNIVERSITY LIBRARY

Table 9.3.2 Selected Reagents to Cleave/Modify and Activate Polymeric Matrix Backbone[a]

Polymer	Reagent	Chemical change produced	Subsequent activation needed	Toxicity of reagents	Cleavage time (hr)	Protein group bonded	Stability of bond	References
Polyacrylamide	Hydrazine	Amide to hydrazide	Nitrous acid yields azide	High	1-5	Amine	Good	Inman and Dintzis (1969)
	Acid pH	Amide to carboxyl	Water-soluble carbodiimide	Low	1-20	Amine	Good	Inman and Dintzis (1969)
Cellulose, agarose	Periodate	Diol to aldehyde	For stability, the imine is reduced by borohydride	Low	1-20	Amine	Good if reduced	Parikh et al. (1974)
Agarose	Acid pH	Anhydrogalactose yields free aldehyde	For stability, the imine is reduced by borohydride	Low	3	Amine	Good if reduced	Stults et al. (1983)
Polyester	Acid pH	Ester to carboxylic acid + alcohol	Water soluble carbodiimide, Woodward's reagent K	Low	1-5	Amine	Good	Rozprimova et al. (1978)
Polyethylene	Concentrated nitric acid	Oxidizes CH_2 to COOH	Water soluble carbodiimide, Woodward's reagent K	Moderate	Variable	Amine	Good	Ngo et al. (1979)
Polystyrene	Concentrated nitric acid	Nitrate aromatic ring	Reduce with Zn/HCl, then diazotize	Moderate	Variable	Histidine tyrosine	Fair	Grubhofer and Schleith (1953)
Nylon	Hydrazine	Amide to hydrazide	Nitrous acid yields azide	High	Variable	Amine	Good	Hornby and Goldstein (1976)
	Triethyloxonium tetrafluoroborate	*O*-Alkylation	None	High?	0.2-0.5	Amine	Fair	Hornby and Goldstein (1976)

[a]Reproduced from Scouten (1987) with permission of Academic Press.

BCU/ASI/CENTRAL UNIVERSITY LIBRARY

employed and it should be opened in the hood. Because of the problems involved in handling CNBr, it is best to dissolve a complete bottle in a defined amount of anhydrous acetonitrile. Short-term storage of this solution (several weeks) at 0°C is possible, but prolonged storage is not recommended.

Materials

Agarose gel (e.g., Sepharose CL-4B)
2 M and 5 M potassium phosphate, pH 12.1
2 M sodium carbonate
2 g/ml CNBr in anhydrous acetonitrile (see recipe)
1 M NaOH
0.2 M sodium bicarbonate, pH 9.0
Sample to be coupled: typically 1 to 10 mg/ml protein, 10 to 200 mg/ml small ligand, or 0.1 to 10 mg/ml nucleic acid
0.1 M potassium phosphate, pH 7.0
Solvent- and CNBr-resistant gloves

1. Wash the agarose gel well with deionized water to remove any preservatives present, then exchange it into 2 M potassium phosphate, pH 12.1, or 2 M sodium carbonate by washing it with 5 vol of the buffer.

   Both buffers work equally well (March et al., 1974). Normally, after washing, the agarose gel is suction filtered with suction maintained until the surface of the agarose gel just barely "cracks." This agarose gel is weighed to obtain the "wet weight" measurement. Although this is a technique-driven procedure and the results vary from person to person, it is the most convenient and widely used method of measuring agarose gel. Other methods, e.g., "dry weight," may be more accurate, but are highly impractical. In most instances, the agarose gel weight is not a critical factor and the person-to-person variation in "wet weight" is not significant.

2. Working in the hood, add 10 g wet weight washed agarose gel at 0°C to 10 ml of ice-cold 5 M potassium phosphate, pH 12.1, or 2 M sodium carbonate, depending on which was used in step 1. Slowly add 1 ml of a 2 g/ml solution of CNBr in anhydrous acetonitrile, with constant stirring and monitoring of the temperature.

   The temperature of the reaction should not exceed 10°C. This can be controlled by the rate of addition of the CNBr solution, by the use of an ice-ethanol bath around the reactor vessel, and/or by adding pieces of ice to the reaction mixture as needed.

   The amount of activation can be varied by increasing or decreasing the volume of CNBr solution. It is preferable to do several test experiments with the total system than to rely on formulas to predict activation levels, even though several such formulas exist (Porath, 1974).

   The author usually uses the weight given by the manufacturer to make up the solution, since any resulting inaccuracy is relatively unimportant. Because the reaction is much more dependent on small changes in temperature, pH, addition rate, etc., than on the precise CNBr concentration, weighing the CNBr is not worth the danger and time involved. All subsequent reactions using CNBr must be done in the hood and wearing appropriate gloves. Excess reagent should be destroyed carefully according to the Material Safety Data Sheet instructions.

3. After the required amount of CNBr is added, keep the mixture under 5°C with stirring for 5 min. Pour the mixture into a suction filtrate apparatus in the hood and filter the agarose gel from the excess reaction mixture.

4. Immediately wash the agarose gel with 5 to 10 vol ice-cold water.

   Destroy the contents of the filter flask, which contains excess CNBr, by slowly adding—in the hood—a 5-fold molar excess relative to the total initial CNBr used, of

PCU IASI / CENTRAL UNIVERSITY LIBRARY

1 M NaOH. Keep the reaction mixture cold by adding pieces of ice. The mixture should be kept in the hood until all CNBr is destroyed.

5. Add an ice-cold solution of protein (1 to 10 mg/ml), small ligand (10 to 200 mg/ml), or nucleic acid (0.1 to 10 mg/ml) in 10 ml of 0.2 M sodium bicarbonate buffer, pH 9.0, to the activated agarose gel.

   Although it is possible to store CNBr-activated agarose gel and activated agarose gel in dehydrated form (i.e., the commercially available CNBr-activated agarose gel) the only practical plan in the laboratory is to couple it to proteins or other biomolecules immediately after preparation (Kohn and Wilchek, 1983b).

6. Place the agarose gel/ligand/buffer solution in a sealed bottle and rotate slowly, end over end, overnight. Collect the agarose gel by suction filtration and wash with 100 ml of 0.1 M potassium phosphate buffer, pH 7.0.

   The temperature depends on ligand lability; for small molecules it is usually 25°C; for proteins it is 4°C. The agarose gel/ligand can be stored, but conditions and length of storage depend on the ligand.

ACTIVATION WITH *p*-NITROPHENYL CHLOROFORMATE

ALTERNATE PROTOCOL 1

Because some materials coupled to CNBr-activated agarose gel lack stability, other activation methods are frequently employed. This alternate protocol describes the use of *p*-nitrophenyl chloroformate in activating agarose gel to produce a stable activated agarose gel that, when coupled to protein or nucleophilic ligands, produces reasonably strong, nonionic coupling bonds to the bound ligand.

***Additional Materials** (also see Basic Protocol)*

30:70 and 70:30 (v/v) acetone/water
Acetone, reagent grade, room temperature and ice cold
Anhydrous acetone (dried over 4 Å molecular sieves)
Anhydrous acetonitrile, room temperature and 4°C (see recipe)
p-Nitrophenyl chloroformate, 800 mg in 12 ml anhydrous acetonitrile
Dimethylaminopyridine
5% acetic acid in dioxane, ice cold
Methanol, ice cold
Anhydrous isopropanol, ice cold
0.5 and 0.1 M potassium phosphate, pH 7.5
0.1 M ethanolamine, pH 7.5

Prepare the agarose gel

1. Wash the agarose gel (Sepharose CL-4b, 10 g wet weight) on a sintered glass funnel successively with 100 ml of each of the following ice-cold solutions:

   30:70 (v/v) acetone/water
   70:30 (v/v) acetone/water
   100% reagent-grade acetone
   anhydrous acetone
   anhydrous acetonitrile.

2. Suspend the agarose gel in 12 ml of 4°C anhydrous acetonitrile containing 800 mg *p*-nitrophenyl chloroformate. Slowly add 3 equivalents of dimethylaminopyridine in anhydrous acetonitrile and stir the mixture for 1 hr at 0° to 4°C.

BOURNS / CENTRAL UNIVERSITY LIBRARY

3. Wash the agarose gel beads successively with 100 ml of each of the following ice-cold solutions:

   reagent-grade acetone
   5% acetic acid in dioxane
   methanol
   anhydrous isopropanol.

4. Store the agarose gel beads at 0° to 4°C in anhydrous isopropanol until used.

   If kept dry, the agarose gel beads may be stored in anhydrous isopropanol for 6 to 12 months.

Perform the coupling reaction

5. On a suction funnel, wash the stored agarose gel (10 g) with ice-cold distilled water to remove all isopropanol.

6. Transfer the agarose gel to 10 ml of 0.1 M potassium phosphate buffer, pH 7.5, containing the ligand at 1 to 10 mg/ml protein, 0.1 to 10 mg/ml nucleic acid, or 10 to 50 mg/ml low-molecular-weight nucleophilic ligand.

7. Place the agarose gel in a bottle, cap the bottle, and tumble it end over end at 4°C for 24 to 48 hr.

   With stable ligands, 25°C for 16 hr will probably be adequate.

8. Wash the agarose gel by suction filtration with 100 ml of 0.1 M phosphate buffer, pH 7.5.

9. Transfer the agarose gel to 100 ml of 0.1 M ethanolamine, pH 7.5.

   For stable low-molecular-weight ligands, solvents with higher pH values, e.g., pure ethanolamine, may be used.

10. Tumble overnight as in step 7.

11. Wash the agarose gel with 0.1 M potassium phosphate, pH 7.5.

12. Add a small sample of the agarose gel to 0.1 M NaOH. If an intense bright yellow color develops, repeat steps 9 to 11.

13. Assay the immobilized ligand as appropriate for its activity.

ALTERNATE PROTOCOL 2

ACTIVATION WITH TRESYL CHLORIDE

This alternate protocol describes the use of tresyl chloride to activate and couple proteins and other nucleophilic ligands to beaded agarose gel. This method is often employed to secure the ligand to the matrix by an exceptionally stable bond. It also produces an activated intermediate that can be stored for at least 6 months in 1 mM HCl at 0°C.

***Additional Materials** (also see Basic Protocol)*

Pyridine (stored over NaOH pellets; see recipe)
Tresyl chloride
5 mM and 1 mM HCl
Protein ligand sample: typically 1 to 10 mg/ml of protein or 10 to 50 mg/ml low-molecular-weight nucleophilic ligand
0.2 M potassium phosphate, pH 8.0
1 M glycine, pH 8.0

RCU IASE CENTRAL UNIVERSITY LIBRARY

1. Wash 10 g (wet weight) agarose gel successively with 100 ml of each of the following ice-cold solutions:

   30:70 (v/v) acetone/water
   70:30 (v/v) acetone/water
   100% reagent-grade acetone
   anhydrous acetone
   anhydrous acetonitrile.

   Although totally anhydrous conditions are not essential, the use of anhydrous solvents is necessary for maximum activation.

2. Add the agarose gel to a bottle containing 5 ml anhydrous acetone and 0.2 ml pyridine (stored over NaOH pellets) and cool the mixture to 0°C in an ice bath.

3. Add 0.25 ml tresyl chloride to the mixture slowly with stirring.

4. After 10 min, wash the agarose gel with 100 ml of each of the following ice-cold solutions given in step 1, but adding 5 mM HCl to the aqueous component of the acetone/water mixtures:

   anhydrous acetonitrile
   anhydrous acetone
   100% reagent-grade acetone
   70:30 (v/v) acetone/water
   30:70 (v/v) acetone water.

5. Wash the agarose gel with 100 ml ice-cold 1 mM HCl and store the agarose gel beads at 4°C in the same dilute HCl solution.

6. Dissolve the ligand to be bound in 10 ml of 0.2 M potassium phosphate buffer, pH 8.0.

   Usually from 1 to 10 mg/ml protein or 10 to 50 mg/ml low-molecular-weight nucleophilic ligand is employed.

7. Rapidly rinse the activated agarose gel with 100 ml ice-cold 0.2 M phosphate, pH 8.0.

8. Mix the washed agarose gel beads with the ligand solution.

   If a protein or temperature-sensitive ligand is employed, rotate the mixture at 4°C overnight. For more robust ligands and/or for higher yields, higher temperatures (25°C) can be employed.

9. Wash the coupled agarose gel beads with 0.2 M potassium phosphate buffer, pH 8.0, at the coupling temperature, and deactivate any residual tresyl groups by incubating the agarose gel overnight with 1 M glycine, pH 8.0.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

***Cyanogen bromide** (Kohn and Wilchek, 1983b)*

CNBr is commercially available as a 5.0 M (~530 mg/ml) solution in acetonitrile (Aldrich). Because CNBr is volatile and very hazardous, *the premade solution is preferred.* The Basic Protocol can be adjusted by using 3.8 ml of the 5 M solution in place of each milliliter of 2 g/ml CNBr needed. If, however, solid CNBr is purchased, it is most convenient and usually safest to make an appropriate (2 g/ml) solution in anhydrous acetonitrile.

continued

BOSTON / CENTRAL UNIVERSITY LIBRARY

For example, a 25-g bottle of *fresh* CNBr may be cooled in a refrigerator overnight (stored upside down, since CNBr lyophilizes and crystallizes on the inside of the lid if stored refrigerated right side up), and then transferred to a hood where it is carefully opened and 50 ml of anhydrous acetonitrile at 0° to 4°C are added. The mixture is stirred with a magnetic stir bar until all of the CNBr is dissolved. The solution should be used within 2 to 4 weeks of preparation and kept at 0° to 4°C until used. As hydrogen cyanide can develop from the reaction of CNBr with moisture, it should *never* be stored in a cold room.

NOTE: *Because CNBr is very hazardous, all precautions noted in the Material Safety Data Sheet (MSDS) should be followed precisely and disposal should be done in accordance with the recommendations! The MSDS should be carefully read before handling CNBr. Old bottles should not be used, nor should CNBr be stored for a prolonged time, as it is reported that age will cause brown discoloration and the formation of explosive degradation products (see Kohn and Wilchek, 1983b).*

Anhydrous solvents

Anhydrous solvents can be prepared by a variety of methods, the most thorough of which are described by Vogel (1978) and Fieser and Fieser (1967). For our purposes, however, the following procedures are usually adequate.

Acetonitrile. Using the commercial anhydrous solvent, to which 4 Å molecular sieves are added for storage, is probably the best procedure. A new bottle of anhydrous acetonitrile is opened and enough 4 Å molecular sieves are added to cover the bottom of the bottle to a depth of ¼ to ½ in.

Pyridine. The same procedure as for acetonitrile is applied, using a new bottle of anhydrous-grade pyridine and adding NaOH pellets instead of molecular sieves. If reagent-grade pyridine is used, NaOH pellets are added and the bottle is allowed to stand at 25°C for 2 to 3 days with occasional shaking (once or twice a day) before using.

Acetone. Acetone is the most difficult solvent to dry and, no matter how dried, cannot be expected to be totally anhydrous. One procedure is to add ½ to 1 in. of calcium sulfate (*not* calcium chloride) to the bottom of 500 ml of acetone in a reagent bottle. The acetone is mixed occasionally by shaking and kept for 4 to 8 hr. It is then filtered and stored in a reagent bottle with ½ to 1 in. of 4 Å molecular sieves. This is not totally anhydrous, but works for these procedures. Consult Vogel (1978) if more stringent conditions are desired.

COMMENTARY

Background Information

Affinity chromatography has benefited both basic research and biotechnology. Without this tool, purification of restriction enzymes, antibodies, and polymerases, for example, would have been much more difficult. Affinity chromatography is a very diverse method of protein purification, encompassing many combinations of matrices, ligands, methods of activation, and coupling (Turkova, 1993).

The basic considerations in designing an affinity chromatography system include the final application of the product, various cost and safety considerations, and potential scale-up concerns. These must all be evaluated when affinity systems are initially planned. The application of the product dictates the amount of nonspecific binding that can be tolerated. Cost and safety considerations determine whether the matrix can have "acceptable" leakage levels and, therefore, whether inexpensive and less durable or expensive and therapeutically safe materials should be used. Scale-up considerations determine the porosity and mechanical stability required of the matrix.

Matrix considerations

The ideal matrix has sufficient porosity to maintain flow, mechanical and chemical stability, high capacity, and very low nonspecific absorption. Such an ideal matrix does not yet exist. The most commonly used matrices are

RCU IASI - CENTRAL UNIVERSITY LIBRARY

naturally occurring hydroxylic polymers, such as agarose, cellulose, and cross-linked dextrans. Agarose gel, the most widely used matrix, is well suited to laboratory benchtop experiments in which small columns are employed, and in which many matrices containing different activation levels, ligands, and functions are often used. For scale-up, however, synthetic organic polymers may be preferred. A new material on the market, Emphaze (3M Company; Coleman et al., 1990), is a preactivated polymer of a vinyl azolactone. Its preactivation makes it very convenient to use and it exhibits relatively low nonspecific binding and high mechanical stability. Its usable flow rate, however, may be less than that of two-component systems, such as HyperD (BioSepra), Poros (PerSeptive Biosystems; Ugelstad et al., 1983; Afeyen et al., 1990, 1991; Regnier, 1991), and Mono-Q/Mono-S (Pharmacia Biotech; Ugelstad et al., 1983). These matrices have relatively fast throughput rates due to the fact that they are multicomponent materials with large pores but reasonably high capacity. With the "gel-in-a-shell" HyperD system, shown in Figure 9.3.2, the synthetic base provides the necessary mechanical stability and the coating of dextran provides an open hydroxylic framework, which minimizes nonspecific binding and maximizes capacity.

Inorganic materials such as silica have often been successfully used in affinity systems when HPLC affinity chromatography is employed. Unfortunately, for large-scale purifications of proteins with a high degree of purity, the nonspecific binding exhibited by virgin silica may be a problem.

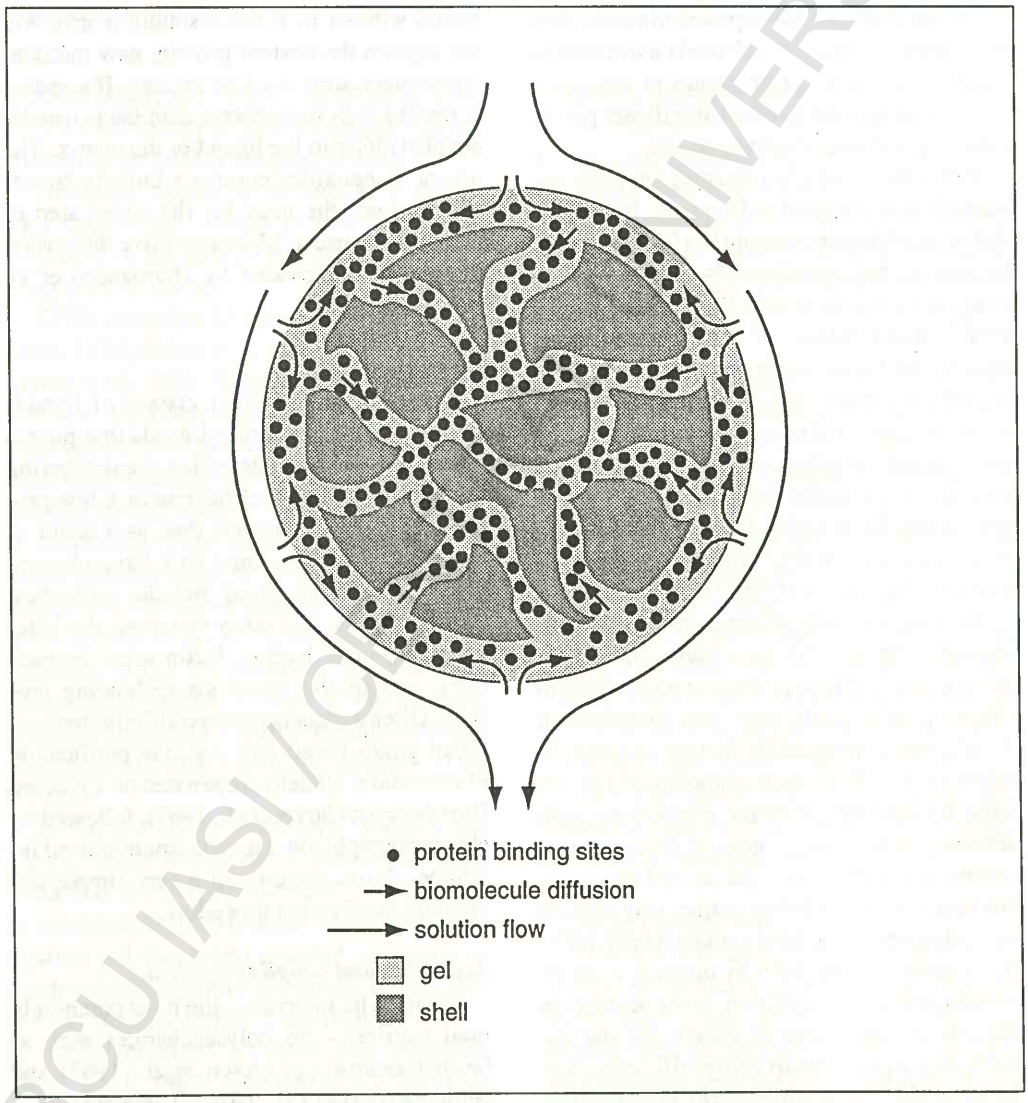


Figure 9.3.2 Rigid HyperD media "gel-in-a-shell." Illustration courtesy of BioSepra, Inc. HyperD is a trademark of BioSepra, Inc.

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IASI / OCR UNIVERSITY LIBRARY

Choice of ligand

Biological ligands used in affinity chromatography can be macromolecular, e.g., proteins and nucleic acids, or low-molecular-weight biomolecular, e.g., substrates, inhibitors, cofactors, antigenic peptides, "phage display" peptides, and biomimetic dyes. Each of these has a biologically relevant interaction with the molecules being purified. Purification based on such immobilized ligands is termed "bioselective adsorption," whereas purification based on resins that bind to a desired molecule via chemical rather than biological interaction is termed "chemiselective adsorption" (Scouten, 1981). Examples of the latter include ion exchange, hydrophobic, boronate, and active thiol ("covalent chromatography") functions bound to a solid matrix. The choice of ligand depends upon which interaction is most nearly unique to the molecule to be purified relative to the other biomolecules in the sample. For example, if it is one of the few glycoproteins in a mixture, a boronate column—which binds selectively to vicinol diols, e.g., sugar groups of glycoproteins—can be used to effect significant purification by a chemiselective process.

Boronates bind glycoproteins via weak covalent bonds (Bergold and Scouten, 1983) such that an equilibrium constant has little meaning. In contrast, biological ligands go from weak to strong noncovalent bonds; they may interact so weakly that a column of immobilized ligands retards the biomolecule being purified only slightly, or, conversely, it may bind almost covalently, as is the case with the biotin-avidin pair. Ligands at either end of the spectrum are unlikely to be useful for normal purification procedures. To be useful, the ligand should bind reasonably selectively, with equilibrium constants K_a from 10^4 to 10^8 M^{-1} (Turkova, 1993).

Although affinity chromatography is considered to be specific for a particular ligand-protein pair, and thus produces remarkable purification in a single step (and sometimes it does!), many nonspecific factors decrease its effectiveness. First, most immobilized ligands have biologically relevant interactions with several proteins—e.g., glucose binds several lectins, many enzymes, and several receptors. Because a given cell-free extract may contain several proteins that bind a single ligand, affinity systems are unlikely to produce a single protein product. In addition, other proteins in the mixture may have an affinity for the immobilized ligand due to nonspecific attraction to the matrix, to the spacer arm, or to charged or hydrophobic components of the ligand. Such nonspecific interactions need to be carefully considered, and minimized, in the design of the affinity media, either by trial and error or by using a matrix known for having little nonspecific binding. Finally, the ligands must be immobilized such that the protein-ligand interaction can still occur after immobilization, e.g., by immobilizing the ligand by means of a nucleophilic residue distant from the point of ligand-protein interaction. Often this means designing special ligands with a nucleophilic attachment site at an appropriate location on the ligand.

Spacers

In many cases it is essential to have a "spacer" between the matrix and the ligand. The spacer is added to eliminate steric hindrance by the matrix, which prevents the ligand from binding the enzyme (see Fig. 9.3.3). To determine if a spacer is needed, immobilize the ligand without it. If the resulting matrix will not capture the desired protein, new matrices with spacer arms must be created. If a spacer is needed, a 2- to 6-carbon diamine is usually coupled either to the ligand or the matrix. The use of bisepoxides creates a built-in spacer and obviates the need for the added step of preparing a spacer. More extensive discussion of spacers is provided by Hermanson et al. (1992).

Ligand classes

There are two distinct classes of ligands (Turkova, 1993): specific ligands that have a reasonably unique interaction (remembering the caveat cited above) for one or a few proteins, and general ligands that, as a result of biological functions, bind to a large class of protein. The former group includes antibodies, most receptors, and many enzymes; the latter group includes lectins, lectin-sugar interactions, and similar broad group–binding proteins. Often a sequence of two affinity steps—a broad group–based one, e.g., the purification of essentially all dehydrogenases on Cibacron Blue (*UNIT 9.2*; Clovis et al., 1987), followed by chromatography on a specific immobilized inhibitor—forms the core of a very simple and efficient purification procedure.

Activation and coupling methods

Hydroxylic matrices—the most commonly used matrices—are polysaccharides such as beaded agarose gel (Axen et al., 1967) and cellulose (Peška et al., 1976). These are generally chemically stable and exhibit minimal non-

BCU IASLIC CENTRAL UNIVERSITY LIBRARY

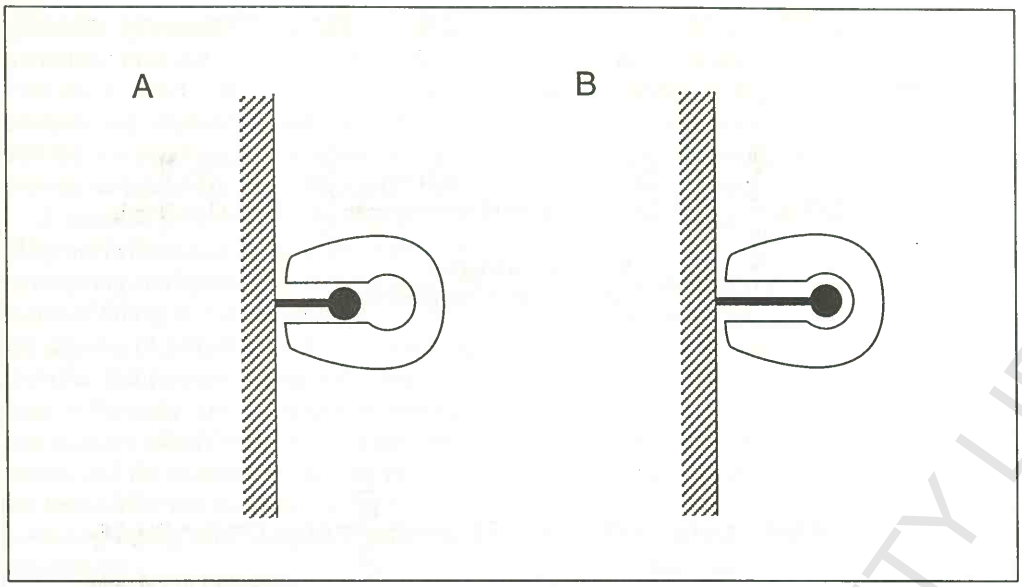


Figure 9.3.3 Effect of spacer arms on enzyme binding to immobilized ligands. (**A**) Direct attachment—the enzyme active site is prevented from reaching the substrate. (**B**) Attachment through a spacer arm—the enzyme active site can freely interact with the substrate. Reproduced from Scouten (1983).

specific adsorption. Most importantly, however, a wide variety of ligands, usually nucleophiles, can be readily coupled to them.

Attaching ligands to hydroxylic matrices, e.g., agarose gel, occurs in two steps: first, activation of the matrix, and second, the coupling process per se.

CNBr activation (Axen et al., 1967; Cuatrecasas, 1970; Porath et al., 1973; Porath, 1974; March et al., 1974; Kohn and Wilchek, 1982, 1983), as described in the Basic Protocol, is the most widely used method, primarily because it is the most studied system. The ligand is readily coupled by this method, but, unfortunately, it is undesirable from the point of view of the safety hazards involved in the activation procedure and the lack of stability of the resulting bond (Parikh et al., 1974; National Research Council, 1981). Furthermore, the procedure generates a positive isourea in the coupling process, which may impart some ion-exchange properties to the resulting product (Fig. 9.3.4A).

p-Nitrophenyl chloroformate (Wilchek and Miron, 1982; Miron and Wilchek, 1987), as described in Alternate Protocol 1, can be used to activate agarose to produce a stable activated agarose gel that when coupled to protein or nucleophilic ligands produces reasonably strong, nonionic bonds to the ligand.

Tresyl chloride (Nilsson and Mosbach, 1980, 1981, 1987; Scouten et al., 1987), as described in Alternate Protocol 2, can be used to activate agarose, resulting in a reasonably stable, but reactive, ester. This can subsequently be coupled to a ligand by an exceptionally stable bond (Fig. 9.3.4B).

Other coupling procedures for hydroxylic matrices worth noting include carbonyldiimidazole (Hearn, 1987), bisepoxide (Sundberg and Porath, 1974), and imine (Stults et al., 1983) systems.

An additional method of creating an activated matrix with synthetic polymers is to form the polymer by polymerizing preactivated monomers. A recent example of this type of synthetic matrix is the previously discussed vinyl azolactone copolymer marketed as Emphaze (3 M Company and Pierce); other new synthetic matrices include two-component systems such as HyperD (gel-in-a-shell) and the macroporous-microporous material Poros, both of which have excellent mechanical stability and fast flow characteristics (see Matrix Considerations). These, along with Pharmacia Biotech's hardened agarose, Superose and Sepharose Fast Flow, are being used in industrial affinity chromatography systems.

Critical Parameters

As noted in Reagents and Solutions, CNBr is very toxic and hazardous. The Material Safety Data Sheet (MSDS) should be read carefully and all precautions should be carefully followed. Note the following in particular (also see Reagents and Solutions):

BRYANT (CENTRAL) UNIVERSITY LIBRARY

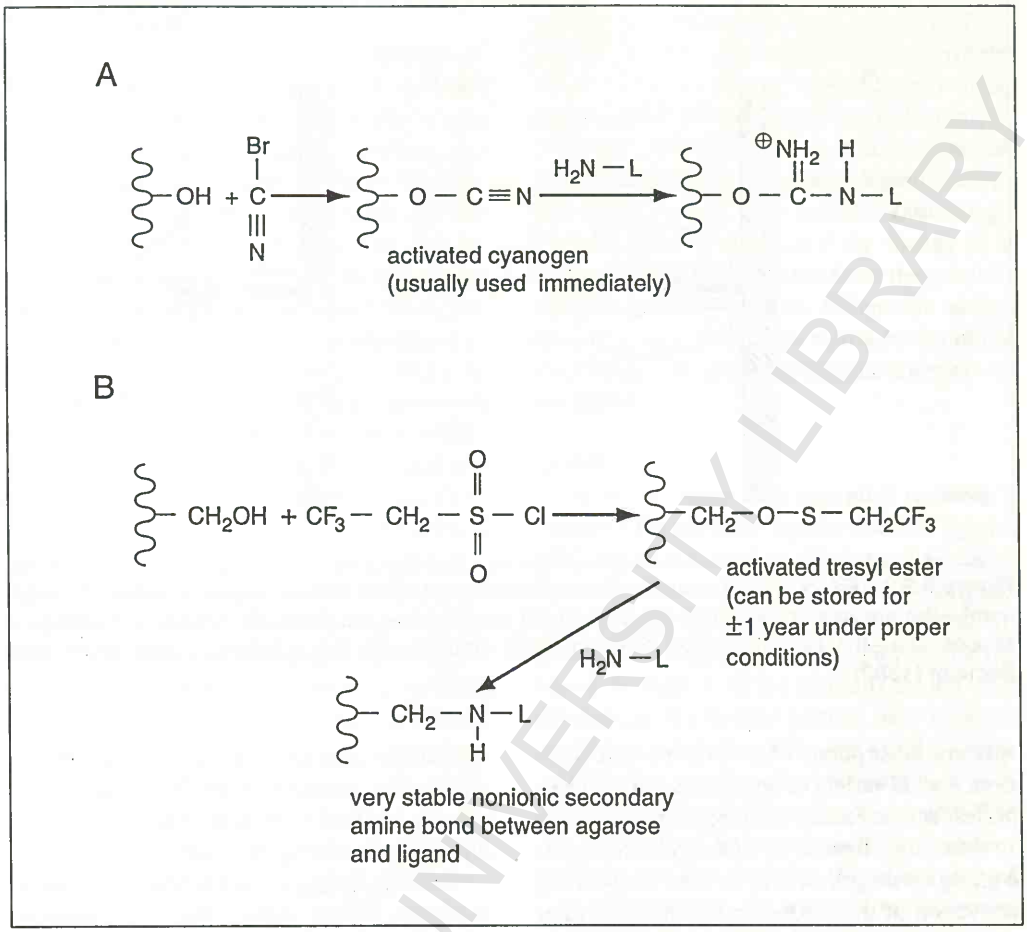


Figure 9.3.4 Activation reactions. (**A**) CNBr activation; (**B**) tresyl chloride coupling.

1. CNBr is very toxic and volatile and releases toxic cyanide fumes.
2. Because CNBr sublimes, bottles should be stored (briefly) in the refrigerator upside down.
3. CNBr should be used only in the hood, with proper equipment, per MSDS instructions, and never in a cold room.
4. Excess CNBr should be destroyed per the MSDS.
5. CNBr should not be stored for prolonged periods.
6. Use of solutions of CNBr in anhydrous acetonitrile is the most convenient method; use of prepared solutions (Aldrich) avoids some of the problems inherent in handling CNBr.

Because of the many variables that influence coupling/activation yields, neither activation nor coupling of ligands to solid matrices is precisely reproducible. For maximum reproducibility, the temperature, rate of addition and mixing, pH (where appropriate), moisture (in anhydrous solvents), and reagent and/or ligand concentration must all be carefully controlled, which is not easy to do. Normally a range of activation and coupling efficiencies is observed, largely because hydrolysis of the activated matrix is a significant process that competes with activation and coupling and that can, sometimes, even be the major reaction. Moreover, in coupling, the reactivity of nucleophilic residues of apparently similar ligands may be quite different, leading to dissimiliar coupling rates and efficiencies. In spite of this, sufficient control of the conditions is possible to achieve reasonably reproducible results *for a given system*, although results and the degree of reproducibility will vary from system to system.

Troubleshooting

It is wise to assay the final degree of activation when using the system described in the Alternate Protocols. This can be done qualitatively for *p*-nitrophenyl chloroformate (Alternate Protocol 1) by adding a small sample of the agarose gel to 0.1 M NaOH. If activation is successful, assuming the washes in step 3 of Alternate Protocol 1 were carried out as described, a bright yellow color will develop. Activation in either Alternate Protocol can be

RCU IASHICE TMAL UNIVERSITY LIBRARY

quantified by elemental analyses for nitrogen (Alternate Protocol 1) or fluoride (Alternate Protocol 2). For CNBr activation, methods of analysis are available (Kohn and Wilchek, 1983b), but usually it is best to base activation success on ligand coupling efficiency.

A common problem in activation and coupling lies in the competition between activation or coupling and hydrolysis. This is usually the result of failing to control temperature, pH, or the dryness of anhydrous solvents, where applicable. Temperature during activation can be kept sufficiently low if mixing is thorough, reactants are added slowly (monitoring temperature), and the reaction vessels are held in an ice bath. CNBr activation mixtures can also be cooled by adding ice whenever the temperature exceeds 5°C.

Anticipated Results

Perhaps the best way to gain an appreciation of the range of potential activation and coupling levels is to check the catalogs of major suppliers (e.g., Pierce, Sigma, and Pharmacia Biotech), where, for example, tresyl chloride is reported in terms of its ability to couple chymotrypsin (4 to 8 mg/ml of 4% cross-linked beaded agarose gel), *p*-nitrophenyl chloroformate is reported in micromoles of activation (10 to 30 μmol/ml of 4% cross-linked beaded agarose gel), and amino acids are immobilized on CNBr-activated 4% beaded agarose gel at from 1 μmol/ml (cysteine) to 20 μmol/ml (histidine). Such broad ranges represent, primarily, the extremes of batch-to-batch and ligand-to-ligand variations. Unfortunately, these values are based sometimes on wet weight and sometimes on dry weight agarose gel; thus, care needs to be taken in interpreting reported activation and/or coupling yields. Moreover, the matrix has a significant effect on yield. For example, Sepharose 4B is a 4% agarose gel with large pore diameters and small surface area, whereas Sepharose 6B is a 6% agarose gel with small pore diameters and large surface area. At a given activation level, Sepharose 6B would couple more of a small ligand, and Sepharose 4B, with its larger pores, might couple more of a very large ligand, e.g., a large protein or nucleic acid. In addition, cross-linked agarose gels (e.g., Sepharose CL-4B; CL = cross-linked) have smaller pores than non-cross-linked agarose gels.

Given the many variables—pore size, surface area, ligand size, ligand reactivity, and lability—it is difficult to predict potential yields for established systems, let alone for systems involving the immobilization of ligands that have never been immobilized. The best suggestion is to peruse catalogs and the literature for *systems* (e.g., agarose gel type and ligand type) that most closely approximate the one to be investigated and to use such a system as the best "predictor" of potential results.

Time Considerations

For the Basic Protocol, activation takes approximately 10 min, washing 10 to 20 min, and coupling is overnight, followed by washing 10 to 15 min. For Alternate Protocol 1, washing the agarose gel takes 20 to 30 min, activation takes 1 hr, and washing requires 30 to 40 min. For coupling, a 5-min wash is followed by a 24- to 48-hr incubation, a 10-min wash, and an overnight deactivation. This is followed by a 10- to 20-min wash and a 5-min test for completeness of deactivation. For Alternate Protocol 2, washing the agarose gel takes 20 to 30 min, and activation requires 10 to 15 min, followed by 30 to 40 min of washing. For coupling, 2 to 3 min is needed for the rapid wash, after which the ligand solution is added to the agarose beads, mixed (2 to 3 min) and the mixture is incubated overnight. This is followed by washing, which takes 20 to 30 min, with a subsequent overnight deactivation step.

Literature Cited

Afeyen, N.B., Gorden, N.F., Maszaroff, I., Varady, L. Fulton, S.P., Yang, Y.B., and Regnier, F.E. 1990. Flow-through particles for the high-performance liquid chromatographic separation of biomolecules: Perfusion chromatography. *J. Chromatogr.* 519:1-29.

Afeyen, N.B., Fulton, S.P., and Regnier, F.E. 1991. Perfusion chromatography packing materials for protein and peptides. *J. Chromatogr.* 544:257-279.

Axen, R., Porath, J., and Ernback, S. 1967. Chemical coupling of peptides and proteins to polysaccharides by means of cyanogen halides. *Nature* 214:1302-1304.

Bergold, A. and Scouten, W.H. 1983. Boronate chromatography. *In* Solid Phase Biochemistry: Analytical and Synthetic Aspects (W.H. Scouten, ed.) pp. 149-188. John Wiley & Sons, New York.

Bethell, G.S., Ayers, J.S., Hancock, W.S., and Hearn, M.T.W. 1979. A novel method of activation of cross-linked agaroses with 1,1′-carbonyldiimidazole which gives a matrix for affinity chromatography devoid of additional charged groups. *J. Biol. Chem.* 254:2572-2574.

BioProbe International. 1986. U.S. Patent 4,582,875.

Clovis, Y.D., Atkinson, A., Bruton, C.J., and Lowe, C.R. 1987. Reactive Dyes in Protein and Enzyme Technology. Stockton Press, New York.

POLITECNICA UNIVERSITY LIBRARY

Coleman, P.L., Walker, M.M., Milbroth, D.S., Stauffer, D.M., Rasmussen, J.K., Krepski, L.R., and Heilmann, S.M. 1990. Immobilization of protein-A at high-density on azolactone—functional polymeric beads and their use in affinity chromatography. *J. Chromatogr.* 512:345-363.

Cuatrecasas, P. 1970. Protein-purification by affinity chromatography. Derivatizations of agarose and polyacrylamide beads. *J. Biol. Chem.* 245:3059-3065.

Domen, P.L., Nevens, J.R., Mallia, A.F., Hermanson, G.T., and Klenk, D.C. 1990. Site-directed immobilization of proteins. *J. Chromatogr.* 510:293-302.

Drobnik, J., Labsky, J., Kudwasarova, H., Saudek, V., and Svec, F. 1982. The activation of hydroxy groups of carriers with 4-nitrophenyl and *N*-hydroxysuccinimide chloroformates. *Biotechnol. Bioeng.* 24:487-493.

Fieser, L.F. and Fieser, M. 1967. Reagents for Organic Synthesis. John Wiley & Sons, New York.

Gribnau, T.C.J. 1977. Coupling of effector-molecules to solid supports. Ph.D. Thesis, University of Nijmegen, Nijmegen, The Netherlands.

Grubhofer, N. and Schleith, L. 1953. Modifizierte Ionenaustauscher als spezifische Adsorbentien. *Naturwissenschaften* 40:508.

Hearn, M.T.W. 1987. Carbonyldiimidazole-mediated immobilization of enzymes and affinity ligands. *Methods Enzymol.* 135:102-117.

Hermanson, G.T., Mallia, A.K., and Smith, P.K. 1992. Immobilized Affinity Ligand Techniques. Academic Press, New York.

Hornby, W.E. and Goldstein, L. 1976. Immobilization of enzymes on nylon. *Methods Enzymol.* 44:118-134.

Inman, J.K. and Dintzis, H.M. 1969. The derivatization of cross-linked polyacrylamide beads. Controlled introduction of functional groups for the preparation of special-purpose, biochemical adsorbents. *Biochemistry* 8:4074-4082.

Kohn, J. and Wilchek, M. 1982. Mechanism of activation of Sepharose and Sephadex by cyanogen-bromide. *Enzyme Microb. Technol.* 4:161-163.

Kohn, J. and Wilchek, M. 1983a. Para-nitrophenylcyanate—an efficient, convenient and nonhazardous substitute for cyanogen-bromide as an activating agent for Sepharose. *Appl. Biochem.* 8:277-285.

Kohn, J. and Wilchek, M. 1983b. Activation of polysaccharide resins by CNBr. *In* Solid Phase Biochemistry: Analytical and Synthetic Aspects (W.H. Scouten, ed.) pp. 599-630. John Wiley & Sons, New York.

Kohn, J. and Wilchek, M. 1984. The use of cyanogen bromide and other novel cyanylating agents for the activation of polysaccharide resins. *Appl. Biochem. Biotechnol.* 9:285-305.

Lawson, T.G., Regnier, F.E., and Weith, H.L. 1983. Separation of synthetic oligonucleotides on columns of macro particulate silica coated with crosslinked polyethylene imine. *Anal. Biochem.* 133:85-93.

Lily, M.D. 1976. Enzymes immobilized to cellulose. *Methods Enzymol.* 44:46-53.

March, S.C., Parikh, I., and Cuatrecasas, P. 1974. Cyanogen-bromide activation of agarose for affinity chromatography. *Anal. Biochem.* 60:149-152.

Miron, T. and Wilchek, M. 1987. Immobilization of protein and ligands using chloroformates. *Methods Enzymol.* 135:84-90.

Murayama, K., Shimada, K., and Yamamoto, T. 1978. Modification of immunoglobulin-G using specific reactivity of sugar moiety. *J. Immunochem.* 15:523-528.

National Research Council, Committee on Hazardous Substances in the Laboratory. 1981. Prudent Practices for Handling Hazardous Chemicals in Laboratories. National Academy Press, Washington, D.C.

Ngo, T.T. and Lenhoff, H.M. 1980. Immobilization of enzymes through activated peptide bonds of protein supports. *J. Appl. Biochem.* 2:373-379.

Ngo, T.T., Laidler, K.J., and Yam, C.F. 1979. Kinetics of acetylcholinesterase immobilized on polyethylene tubing. *Can. J. Biochem.* 57:1200-1203.

Nilsson, K. and Mosbach, K. 1980. *p*-Toluenesulfonyl chloride as an activating agent of agarose for the preparation of immobilized affinity ligands and proteins. *Eur. J. Biochem.* 112:397-402.

Nilsson, K. and Mosbach, K. 1981. Immobilization of enzymes and affinity ligands to various hydroxyl group carrying supports using high reactive sulfonyl chlorides. *Biochemistry* 102:449-457.

Nilsson, K. and Mosbach, K. 1987. Tresyl chloride–activated supports for enzyme immobilization. *Methods Enzymol.* 135:65-78.

Parikh, I., March, S., and Cuatrecasas, P. 1974. Topics in the methodology of substitution reactions with agarose. *Methods Enzymol.* 34:70-102.

Peška, J., Štamberg, J., Hradil, J., and Ilavsky, M. 1976. Cellulose in bead form: Properties related to chromatographic uses. *J. Chromatogr.* 125:455-469.

Porath, J. 1974. Preparation of cyanogen bromide-activated agarose gels. Methods Enzymol. 34:13-30.

Porath, J., Aspberg, K., Drevin, H., and Axen, R. 1973. Preparation of cyanogen bromide activated agarose gels. *J. Chromatogr.* 86:53-56.

Regnier, F.E. 1991. Perfusion chromatography. *Nature* 350:634-635.

Rozprimova, L., Franek, F., and Kubanek, V. 1978. Utilization of powder polyester in making insoluble antigens and pure antibodies. *Czech Epidemiol. Mikrobiol. Immunol.* 27:335-341.

RCU IASI CENTRAL UNIVERSITY LIBRARY

Scouten, W.H. 1981. Affinity Chromatography: Bioselective Absorption on Inert Matrices. Wiley-Interscience, New York.

Scouten, W.H. 1983. Solid Phase Biochemistry: Analytical and Synthetic Aspects. John Wiley & Sons, New York.

Scouten, W.H. 1987. Immobilization techniques for enzymes. *Methods Enzymol.* 135:30-65.

Scouten, W.H., Van den Tweel, W., Kromenburg, H., and Dekker, M. 1987. Colored sulfonyl chloride as an activating agent for hydroxylic matrices. *Methods Enzymol.* 135:79-84.

Stults, N.L., Lin, P., Hardy, M., Lee, Y.G., Uchida, Y., Tsukada, Y., and Sugimori, T. 1983. Immobilization of proteins on partially hydrolyzed agarose beads. *Anal. Biochem.* 135:392-400.

Sundberg, L. and Porath, J. 1974. Preparation of adsorbents for biospecific affinity chromatography: Attachment of group-containing ligands to insoluble polymers by means of bifunctional oxiranes. *J. Chromatogr.* 90:87-98.

Turkova, J. 1993. Bioaffinity Chromatography, 2nd ed. Elsevier, Amsterdam.

Ugelstad, J., Söderberg, L., Berge, A., and Bergstrom, J. 1983. Monodisperse polymer particles: A step forward for chromatography. *Nature* 303:95-96.

Vogel, A. 1978. Textbook of Practical Organic Chemistry, 4th ed. Longman, New York.

Voivodov, K., Chan, W.-H., and Scouten, W.H. 1993. Chemical approaches to oriented immobilization. *Makromol. Chem.* 70/71:275-283.

Wilchek, M. and Miron, T. 1982. Immobilizations of enzymes and affinity ligands onto agarose via stable and uncharged carbamate linkages. *Biochem. Int.* 4:629-635.

Key References

Hermanson, et al., 1992. See above.

A very good practical laboratory manual with good description of methods.

Scouten, W.H. 1981. See above.

An older, simple approach with a good mix of theory and practical laboratory methods.

Turkova, J. 1993. See above.

The most recent and comprehensive volume available in the field.

Contributed by William H. Scouten
Utah State University
Logan, Utah

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Metal-Chelate Affinity Chromatography

Recombinant proteins engineered to have six consecutive histidine residues on either the amino or carboxyl terminus can be purified using a resin containing nickel ions (Ni^{2+}) that have been immobilized by covalently attached nitrilotriacetic acid (NTA). This technique, known as metal-chelate affinity chromatography (MCAC), can readily be performed with either native or denatured protein. The Strategic Planning section discusses techniques for creating a fusion protein consisting of the protein of interest with a histidine tail attached (for purification by MCAC; see *UNIT 6.5* for an example purification of HIV-1 integrase). The Basic Protocol describes expression of histidine-tail fusion proteins and their purification in native form by MCAC. Two alternate protocols describe purification of histidine-tail fusion proteins by MCAC under denaturing conditions and their renaturation by either dialysis or solid-phase renaturation. Support protocols are provided for analysis of the purified product and regeneration of the NTA resin. All of these protocols are easily adaptable to any protein expression system.

Prior to large-scale preparation, the cells should be tested for expression of the protein in soluble form (see *UNITS 6.1 & 6.2*). Even if protein is expressed mostly in insoluble form (i.e., bacterial inclusion bodies; *UNIT 6.3*), there may be a small fraction that remains soluble and can therefore be purified using the Basic Protocol. If little or no soluble protein is observed, one of the denaturing protocols (Alternate Protocols 1 and 2) should be used.

NOTE: All solutions and equipment coming into contact with live cells must be sterile.

STRATEGIC PLANNING

MCAC is designed to purify a specific protein whose complementary DNA (cDNA) sequence is available. A protein-expression system is selected and the cDNA is inserted into the appropriate expression vector (see *UNIT 5.1*). The cDNA sequence must encode a minimum of six histidines at either the amino or carboxyl terminus and must include an initiator methionine at the amino terminus and a termination codon at the carboxyl terminus. This can be accomplished by polymerase chain reaction (PCR) using primers

A 5′-GGGNNNNNNATGCATCATCATCATCATCAT . . . $N_{15\text{-}30}$-3′
RE site – MetHisHisHisHisHisHis . . .

B 5′-GGGNNNNNNTTAATGATGATGATGATGATG . . . $N_{15\text{-}30}$-3′
RE site – ENDHisHisHisHisHisHis . . .

Figure 9.4.1 Sequences of primers required to create histidine tails at protein termini. Functions (e.g., protein sequence) are marked below. Three guanines are included at the 5′ ends of each primer to facilitate restriction enzyme digestion of the PCR product prior to subcloning. NNNNNN represents a unique restriction enzyme site compatible with the selected vector. $N_{15\text{-}30}$ represents 15 to 30 additional nucleotides specific to the cDNA beginning with the second codon. Met represents an initiator methionine and END represents a termination codon. (**A**) Sequence of 5′ primer used to create a histidine tail at the amino terminus. The 3′ primer should include a second unique restriction enzyme site and the final 5 to 10 codons (including a stop codon) of the cDNA sequence. (**B**) Sequence of 3′ (antisense) primer used to create a histidine tail at the carboxyl terminus. The 5′ primer should contain a second unique restriction site and the first 5 to 10 codons of the cDNA sequence.

Copyright © 1996 by John Wiley & Sons, Inc.

BOSTON UNIVERSITY LIBRARY

containing unique restriction-site sequences at the 5′ end (Fig. 9.4.1). Properly designed primers will permit insertion and expression of the cDNA in the correct reading frame.

An alternative to PCR is to subclone the cDNA into an existing vector, synthesize complementary oligonucleotides containing the hexahistidine sequence with compatible restriction enzyme site overhangs, and insert them into the subcloned cDNA (Struhl, 1987). It is important to insert the oligonucleotide into the cDNA near one of the termini in the correct reading frame.

Several companies (e.g., Qiagen, Novagen, and Invitrogen; see *SUPPLIERS APPENDIX*) sell expression systems that permit direct subcloning of cDNAs into vectors already containing oligohistidine tail sequences along with adjacent protease cleavage sites to allow removal of the tail after purification. Vectors have been developed for expression of histidine fusion proteins by bacteria, yeast, baculovirus, vaccinia, and various eukaryotic promoters. The protocols that follow are designed for use with a Novagen pET vector (Fig. 9.4.2), but may easily be modified for other vectors.

BASIC PROTOCOL

NATIVE MCAC FOR PURIFICATION OF SOLUBLE HISTIDINE-TAIL FUSION PROTEINS

This protocol describes expression and purification of histidine fusion proteins in *E. coli*. A desired protein is induced in bacteria, which are then harvested and lysed. The lysate is loaded directly onto a column containing Ni^{2+}-NTA resin and the protein is eluted in nearly pure form with MCAC buffer containing imidazole.

Materials

M9ZB medium (see recipe) containing 50 µg/ml ampicillin and 25 µg/ml chloramphenicol
E. coli BL21(DE3)pLysS or other suitable strain (Novagen) containing a pET vector (Fig. 9.4.2) expressing a histidine-tail fusion protein
0.1 M IPTG, filter sterilized
NTA resin slurry: 50% (w/v) suspension in 20% (v/v) ethanol (Qiagen)
100 mM $NiSO_4 \cdot 6H_2O$
MCAC-0, MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, MCAC-200, and MCAC-1000 buffers (see recipe)
150× protease inhibitor cocktail (see recipe)
10% (v/v) Triton X-100
1 M $MgCl_2$
MCAC-EDTA buffer (see recipe)
DNase I solution (see recipe)

Centrifuge with Beckman JA-20 rotor or equivalent
1 × 10–cm glass or polypropylene column

Additional reagents and equipment for growth of bacteria in liquid medium (*UNIT 5.3*) and analysis and processing of purified proteins by SDS-PAGE (see Support Protocol 1)

Express the protein

1. Inoculate 10 ml M9ZB/ampicillin/chloramphenicol with *E. coli* BL21(DE3)pLysS containing a pET vector expressing a desired histidine-tail fusion protein. Grow overnight with shaking at 37°C (see *UNIT 5.3*).
2. Inoculate 100 ml M9ZB/ampicillin/chloramphenicol with 1 ml of the overnight culture and grow with shaking at 37°C to OD_{600} = 0.7 to 1.0.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

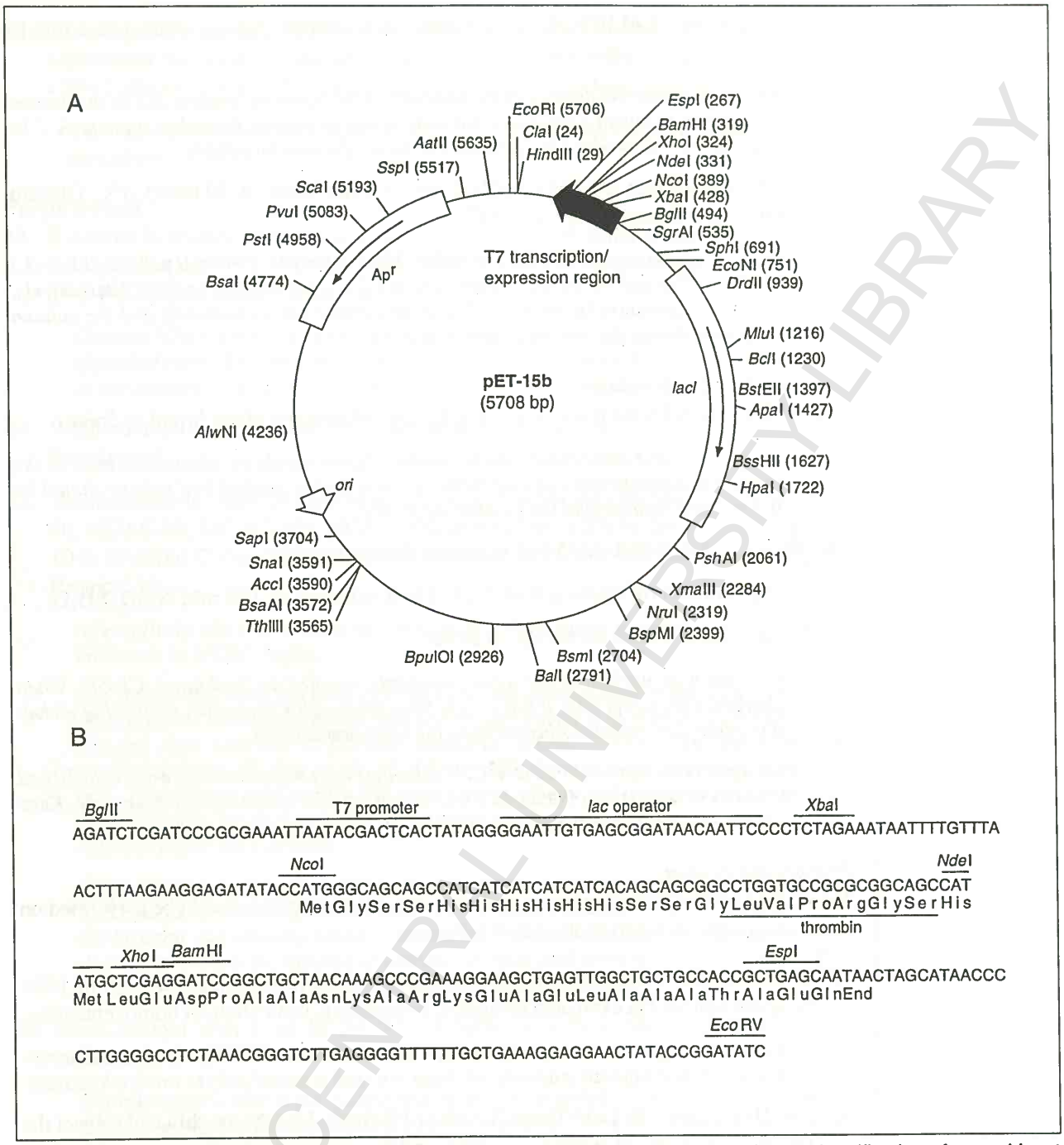


Figure 9.4.2 pET-15b, one of a series of pET vectors designed for cloning, expression, and purification of recombinant proteins. Reprinted with the permission of Novagen. **(A)** pET-15b vector; **(B)** sequence of pET-15b cloning/expression region. The target gene is cloned into the pET plasmid such that its expression is under the control of bacteriophage T7 transcription and translation signals (see *UNIT 5.2*). pET-15b encodes an amino-terminal His·Tag leader that allows purification of the resulting recombinant protein over Ni^{2+}-NTA resin. Following purification, the His·Tag can be removed by thrombin cleavage (see *UNIT 6.5*). The pET series are derivatives of vectors originally described by Studier et al. (1990) and are available from Novagen.

BCU IAS CENTRAL UNIVERSITY LIBRARY

3. Add 1 ml of 0.1 M IPTG (to 1 mM final) and continue shaking incubation 1 to 3 hr at 37°C.

Incubation time will depend on the solubility of the expressed protein. A 3-hr incubation will allow maximal expression but may result in mostly insoluble protein. A 1-hr incubation will produce less protein but more of it may be soluble.

4. Centrifuge 10 min at 4400 × *g* (5000 rpm in a Beckman JA-20 rotor), 4°C. Discard supernatant. Freeze pellet at –70°C.

It is not necessary to wash the cell pellet. The wet weight of the cell pellet will be ~0.5 g. The pellet can be stored indefinitely at –70°C before proceeding. Alternatively, extract preparation (steps 9 to 13) can be carried out immediately and the column prepared during the centrifugation at step 13.

Prepare the affinity column

5. Add 0.2 ml NTA resin slurry to a 1 × 10–cm column and allow liquid to drain.

During this and subsequent column washes, liquid should be allowed to drain to the top of the packed resin bed that forms as resin settles (packed bed volume should be 0.1 ml) and resin should not be allowed to dry.

6. Wash column with 1 ml (5 bed volumes) deionized water.

7. Charge column by washing with 1 ml (5 bed volumes) of 100 mM $NiSO_4 \cdot 6H_2O$.

8. Wash column with 2 ml MCAC-0 buffer.

NTA resin is nitrilotriacetic acid covalently coupled to Sepharose CL-6B. When charged with nickel ions, it has a light blue-green color, and when stripped of nickel, it is white. See supplier's instructions for resin preparation.

Charged resin can be stored at 4°C. If column is to be stored >1 day, wash with 10 bed volumes of 20% ethanol and add 1 bed volume of 20% ethanol prior to storage. Keep column sealed to prevent evaporation.

Prepare the extract

IMPORTANT NOTE: Beginning with this step, all procedures should be performed on ice or in a cold room unless otherwise indicated.

9. Thaw cell pellet (from step 4) on ice. Add 5 ml MCAC-0 buffer and 33 µl of 150× protease inhibitor cocktail and resuspend by pipetting, sonication, or homogenization.

All MCAC buffers contain phenylmethylsulfonyl fluoride (PMSF) as a protease inhibitor; to reduce expense, protease inhibitor cocktail is added only to the crude extract.

10. Add 0.05 ml of 10% (v/v) Triton X-100 (0.1% final). Mix thoroughly and subject the sample to 3 cycles of freezing at –70°C and thawing on ice.

Ionic detergents may interfere with binding of the protein to the resin and should not be used.

Cell lysis is evidenced by a visible increase in viscosity. The pLysS plasmid in these cells encodes an endogenous lysozyme that eliminates the need for exogenous lysozyme treatment to disrupt the bacterial cell wall.

11. Add 0.05 ml of 1 M $MgCl_2$ (final concentration 10 mM) and 0.05 ml of DNase I solution (final concentration 10 µg/ml DNase I). Mix gently and incubate 10 min at room temperature.

The DNase I treatment reduces the viscosity of the lysate.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

12. Centrifuge 15 min at 27,000 × *g* (15,000 rpm in JA-20 rotor), 4°C. Decant the supernatant into a clean container on ice and discard the pellet. Set aside and freeze a 10-µl aliquot at –70°C for later analysis by SDS-PAGE (see Support Protocol 1).

The supernatant can be frozen at –70°C indefinitely before continuing with the procedure.

Purify protein

13. If extract is frozen, thaw on ice. Load onto Ni^{2+}-NTA column and allow to flow through at a rate of 10 to 15 ml/hr. Collect column flowthrough and save for SDS-PAGE (see Support Protocol 1).

Charged NTA resin has a capacity of 5 to 10 mg histidine-tagged protein per milliliter of packed resin. The amount of extract that can be loaded on the column will depend on the amount of soluble histidine-tagged protein in the extract.

14. Wash column with 5 ml MCAC-0 buffer at a flow rate of 20 to 30 ml/hr. Discard flowthrough.

15. Wash column in stepwise fashion with 5 ml each of MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, MCAC-200, and MCAC-1000 buffers at a flow rate of 10 to 15 ml/hr. Collect 0.5-ml fractions and save on ice for SDS-PAGE (see Support Protocol 1).

Alternatively, the column can be eluted with a 5-ml linear gradient of 0 to 400 mM imidazole in MCAC buffer.

The second and third fractions of each wash will contain most of the eluted proteins. Most proteins with hexahistidine tails will remain bound in 60 mM imidazole (MCAC-60) and elute with 100 to 200 mM imidazole (MCAC-100 or -200); therefore, the purified protein will elute in MCAC-100 or -200. Proteins with longer histidine tails (e.g., 10 residues) bind to Ni^{2+}-NTA with greater affinity and require higher imidazole concentrations for elution. However, optimum washing and elution conditions must be determined for each protein.

Once optimum washing and elution conditions are established, it is possible to prepare the crude extract in a buffer that contains the highest imidazole concentration in which the histidine tail remains bound to the Ni^{2+}-NTA (e.g., MCAC-40 or -60 buffer). This decreases nonspecific binding of proteins to resin and permits use of a single buffer for extract preparation, column loading, and column washing.

16. Elute column with 1 ml MCAC-EDTA buffer at a flow rate of 10 to 15 ml/hr, collecting 0.5-ml fractions.

The blue-green color of the column will disappear as nickel is removed by EDTA. More tightly bound proteins may be found in these fractions. The resin can now be recharged (repeat steps 6 to 8) and the column reused.

If protein is eluted adequately with imidazole (as determined by overall yield for subsequent preparations), EDTA washing can be omitted. The column can be reequilibrated with MCAC-0 buffer and the purification repeated. The same column can be used three to five times before EDTA stripping and nickel recharging are necessary. Only one protein should be purified on any given column.

17. Analyze fractions for the presence of eluted protein.

An ultraviolet (280-nm) absorbance flow monitor is helpful for following column elution but is not necessary. An alternative is to measure the OD_{280} of individual fractions to identify protein-containing fractions. However, imidazole will also absorb at 280 nm.

A quick and easy method to determine which fractions contain eluted protein is to place 2 µl undiluted Protein Assay Dye Reagent Concentrate (Bio-Rad) on a piece of

BOWLES GREEN UNIVERSITY LIBRARY

Parafilm, add 8 µl from fraction to be tested, and mix by pipetting up and down. Immediate appearance of blue color indicates that the fraction contains protein. This does not work in the presence of Triton X-100 because the detergent itself produces an intense blue color; for this reason, Triton X-100 is excluded from the washing and elution buffers.

18. Combine the fractions containing eluted protein and remove a 10-µl aliquot for SDS-PAGE (see Support Protocol 1). Freeze the remainder in smaller aliquots at –70°C or in liquid nitrogen.

    If a different buffer for the protein is desired (e.g., for proteolytic removal of the histidine tail), the protein should be dialyzed against the buffer of choice to remove the MCAC buffer prior to storage at –70°C or in liquid nitrogen.

19. If time permits, proceed immediately to analysis of fractions by SDS-PAGE and processing of protein (see Support Protocol 1). Otherwise, freeze all samples at –70°C until ready for analysis and processing.

ALTERNATE PROTOCOL 1

DENATURING MCAC FOR PURIFICATION OF INSOLUBLE HISTIDINE-TAIL FUSION PROTEINS

High-level expression of foreign proteins in bacteria and other cells frequently results in poor solubility of the expressed protein (see *UNITS 5.1 & 5.2*). These insoluble proteins form inclusion bodies in bacteria, and strong chaotropic agents such as guanidine, urea, or SDS are usually required to solubilize them. These agents denature the protein and destroy the secondary structure that is essential to other affinity purification methods (e.g., maltose-binding protein or glutathione-*S*-transferase fusion proteins). A significant advantage of metal-chelate affinity chromatography is that the oligohistidine tail will bind to the Ni^{2+}-NTA resin even when the protein is denatured. In denaturing MCAC, the protein extract is solubilized with 6 M guanidine and the entire affinity purification procedure is carried out in guanidine. The purified, denatured protein is renatured during dialysis.

Additional Materials (also see Basic Protocol)

GuMCAC-0, GuMCAC-20, GuMCAC-40, GuMCAC-60, GuMCAC-100, and GuMCAC-500 buffers (see recipe)
GuMCAC-EDTA buffer (see recipe)
Appropriate final buffer for protein (e.g., for proteolytic cleavage or long-term storage)
Guanidine·HCl
Additional reagents and equipment for analysis and processing of purified proteins (see Support Protocol 1) and dialysis (*APPENDIX 3B & UNIT 4.4*)

Express the protein

1. Prepare the pellet of *E. coli* expressing a histidine-tail fusion protein (see Basic Protocol, steps 1 to 4).

   The pellet can be stored indefinitely at –70°C before proceeding. Alternatively, extract preparation (steps 5 and 6) can be carried out immediately and the column prepared during the centrifugation at step 6.

Prepare the affinity column

2. Prepare column (see Basic Protocol, steps 5 to 7).

3. Wash column with 2 ml GuMCAC-0 buffer.

   During this and subsequent column washes, liquid should be allowed to drain to top of packed resin bed and resin should not be allowed to dry.

BCU IASI CENTRAL UNIVERSITY LIBRARY

Prepare cell extract

4. Thaw cell pellet (from step 1) on ice. Resuspend in 5 ml GuMCAC-0 buffer by pipetting, sonication, or homogenization.

5. Freeze 10 min at –70°C and thaw at room temperature.

   Protease inhibitors are omitted because proteases are inactivated by guanidine. Triton X-100 is not needed at this step. Freezing is not necessary but is included because it ensures complete lysis of cells.

   Subsequent steps can be performed at room temperature. However, if solid-phase renaturation is used (see Alternate Protocol 2), it is better to maintain lower temperatures throughout the process.

6. Gently mix samples for 30 min using a rocker, rotating mixer, or magnetic stirrer. Centrifuge 15 min at 27,000 × *g* (15,000 rpm in Beckman JA-20 rotor), 4°C. Decant supernatant into a clean container and discard pellet. Set aside a 10-µl aliquot for analysis by SDS-PAGE (see Support Protocol 1).

   The supernatant can be frozen at –70°C indefinitely before continuing with the procedure.

Purify protein

7. If extract from step 5 is frozen, thaw at room temperature. Load onto Ni^{2+}-NTA column and allow to flow through at a rate of 10 to 15 ml/hr. Collect flowthrough and save a 10-µl aliquot for SDS-PAGE (see Support Protocol 1).

8. Wash column with 5 ml GuMCAC-0 buffer at a rate of 20 to 30 ml/hr. Discard the flowthrough.

9. Wash column in stepwise fashion with 5 ml GuMCAC-20, -40, -60, -100, and -500 buffers at a rate of 10 to 15 ml/hr. Collect 0.5-ml fractions and save for SDS-PAGE (see Support Protocol 1).

   The second and third fractions from each wash will contain most of the unbound protein. The histidine tail binds slightly less avidly under denaturing conditions. Lower imidazole concentrations are therefore required for washing and elution than in the Basic Protocol.

10. Elute with 1 ml GuMCAC-EDTA buffer at a rate of 10 to 15 ml/hr, collecting 0.5-ml fractions.

11. Identify fractions containing the protein, pool together, transfer to dialysis tubing, and seal.

    Alternatively, fractions can be frozen at –70°C indefinitely before continuing with the procedure.

    Guanidine precipitates in the presence of SDS and must be removed by dialysis before SDS-PAGE. An alternative technique employs buffers that switch from 6 M guanidine to 8 M urea during affinity column washing (Stüber et al., 1990). This permits samples to be taken directly from urea fractions without dialysis and analyzed by SDS-PAGE or injected into animals for antibody production.

Renature purified protein by dialysis

12. Prepare appropriate final buffer for protein (e.g., for proteolytic cleavage or long-term storage) and add sufficient guanidine to bring final concentration to 4 M.

13. Dialyze purified protein from step 11 for ≥2 hr at 4°C against 500 ml buffer/4 M guanidine (see *APPENDIX 3B*).

    The MWCO of the dialysis membrane should be chosen to be smaller than the MW of the purified protein. In most cases an MWCO of 12 to 14 kDa is sufficient.

BUDAPEST CENTRAL UNIVERSITY LIBRARY

14. Remove 250 ml buffer/guanidine and add 250 ml buffer without guanidine. Continue dialysis ≥2 hr. Repeat.

    With some proteins, renaturation by dialysis may require longer dialysis periods and more gradual decrements in the guanidine concentration of the buffer. Conditions for each protein must be determined empirically.

15. Remove dialysis bag to a container containing 500 ml of fresh buffer without guanidine at 4°C. Continue dialysis 2 hr to overnight.

16. Remove sample from dialysis bag, divide into aliquots, and freeze at –70°C or in liquid nitrogen.

    If protein precipitates during dialysis, solid-phase renaturation (see Alternate Protocol 2), in which protein bound to the column is renatured before elution, should be employed.

17. Analyze fractions and process protein (see Support Protocol 1).

ALTERNATE PROTOCOL 2

SOLID-PHASE RENATURATION OF MCAC-PURIFIED PROTEINS

In Alternate Protocol 1, removal of denaturants by dialysis will occasionally lead to precipitation of protein, possibly due in part to entanglement or aggregation of separate protein molecules as they refold. To avoid this problem, solid-phase renaturation may be attempted. In this procedure, protein extract is prepared and bound to the column under denaturing conditions. A series of washes removes the denaturing agent before the target protein is eluted and the resulting renatured protein is eluted from the column under native conditions.

***Additional Materials** (also see Basic Protocol)*

1:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)
3:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)
7:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)

1. Prepare protein extract, bind to column, and wash with GuMCAC buffers (see Alternate Protocol 1, steps 1 to 9).

2. Wash column with 5 ml of 1:1 (v/v) MCAC-20/GuMCAC-20 buffer.

   During this and subsequent washes, liquid should be allowed to drain just to top of packed resin bed and resin should not be allowed to dry.

3. Wash column with 5 ml of 3:1 (v/v) MCAC-20/GuMCAC-20 buffer.

4. Wash column with 5 ml of 7:1 (v/v) MCAC-20/GuMCAC-20 buffer.

5. Wash column with MCAC buffers, elute proteins, and analyze (see Basic Protocol, steps 15 to 19).

   Slow elution (between 1 and 2 hr) with a 5-ml linear gradient from 100% GuMCAC-20 buffer to 100% MCAC-20 buffer may also yield efficient renaturation.

BCU IASI CENTRAL UNIVERSITY LIBRARY

ANALYSIS AND PROCESSING OF PURIFIED PROTEINS

The success of the purification scheme (particularly during a small pilot study) should be monitored at each stage by SDS-PAGE. If the fusion protein contains a specific protease cleavage site, the histidine tail can be removed using an appropriate proteolytic procedure, if desired.

Materials

Fractions from MCAC column purification (crude extract, flowthroughs, and purified protein; see Basic Protocol or Alternate Protocols 1 or 2)
2× SDS sample buffer (*UNIT 10.1*)
MCAC-0 buffer (see recipe)

Additional reagents and equipment for one-dimensional SDS-PAGE (*UNIT 10.1*), cleavage of proteins with factor Xa *or* thrombin (*UNIT 6.5*), and dialysis (*UNIT 4.4 & APPENDIX 3B*)

1. Thaw aliquots of fractions to be analyzed on ice. Mix 5 µl from crude extract and crude flowthrough fractions and 10 µl from the second and third fractions from each washing step with an equal volume of 2× SDS sample buffer.

2. Load samples onto a standard SDS-PAGE gel. Run gel and visualize to identify the fractions containing purified protein (*UNIT 10.1*).

   Guanidine must be removed by dialysis prior to addition of SDS sample buffer.

3. Thaw the remaining aliquots of fractions containing purified protein, dialyze against the appropriate proteolysis buffer, and carry out cleavage procedure if desired. If necessary, after cleavage dialyze the protein against an appropriate storage buffer and freeze in aliquots.

   The size of the cleaved histidine tail will generally be <3 kDa, depending on the design of the fusion protein. This fragment will usually be removed by the dialysis. It may also be removed by ultrafiltration (which will also concentrate the protein), size-exclusion chromatography, or a second MCAC (see Basic Protocol, steps 13 to 19). The advantage of a second MCAC is that the histidine tail and any uncleaved protein will bind to the column and only the cleaved protein will be found in the flowthrough from step 12. Aliquots of each fraction should be analyzed by SDS-PAGE to verify recovery of pure protein.

   If an expression vector other than pET has been used, and the fusion protein does not contain an appropriate cleavage site for factor Xa or thrombin, or if the histidine tail will not interfere with subsequent studies, omit the cleavage procedure. Presence of the histidine tail will not generally affect the protein's biologic functions (see Background Information).

B. C. U. „M. EMINESCU" IAŞI
C H I - E

BCU IASI / CENTRAL UNIVERSITY LIBRARY

NTA RESIN REGENERATION

NTA resin can be repeatedly charged with Ni^{2+} and stripped. Over time, however, some protein residue will accumulate and decrease the efficiency of the resin, resulting in slow flow rates or lack of blue-green color after charging with Ni^{2+}. Periodic regeneration (e.g., after 5 to 10 cycles of charging and stripping) permits recycling of the resin for long-term use.

Materials

2.5 ml used NTA resin (packed volume)
Stripping solution: 0.2 M acetic acid/6 M guanidine·HCl
2% (w/v) SDS
20%, 25%, 50%, 75%, and 100% (v/v) ethanol
0.1 M EDTA, pH 8.0

1. Wash resin with 5 ml stripping solution.

   In this and all subsequent washes, liquid added to the column should be allowed to drain to top of packed resin bed and resin should not be allowed to dry.

2. Wash with 5 ml water.
3. Wash with 7.5 ml of 2% (w/v) SDS.
4. Wash with 2.5 ml of 25% (v/v) ethanol.
5. Wash with 2.5 ml of 50% (v/v) ethanol.
6. Wash with 2.5 ml of 75% (v/v) ethanol.
7. Wash with 12.5 ml of 100% (v/v) ethanol.
8. Wash with 2.5 ml of 75% (v/v) ethanol.
9. Wash with 2.5 ml of 50% (v/v) ethanol.
10. Wash with 2.5 ml of 25% (v/v) ethanol.
11. Wash with 2.5 ml water.
12. Wash with 12.5 ml of 0.1 M EDTA, pH 8.0.
13. Wash with 7.5 ml water. Proceed either to nickel charging (see Basic Protocol, step 7) or to step 14 below for long-term storage of the resin.
14. Add 2.5 ml of 20% (v/v) ethanol to resin and store at 4°C.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

DNase I solution

Dissolve lyophilized DNase I (U.S. Biochemical) at a final concentration of 1 mg/ml in 50% glycerol/1 mM $CaCl_2$. Store ≤1 year at –20°C.

GuMCAC buffers (store ≤6 months at 4°C)

GuMCAC-0 buffer:
20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl

GuMCAC-500 buffer:
20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl
0.5 M imidazole

GuMCAC-20, GuMCAC-40, GuMCAC-60, and GuMCAC-100 buffers:

These buffers (containing different concentrations of imidazole) are made by mixing GuMCAC-0 buffer and GuMCAC-500 buffer in the appropriate ratios: e.g., for GuMCAC-20 buffer, use 96:4 (v/v) GuMCAC-0/GuMCAC-500.

GuMCAC-EDTA buffer

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl
0.1 M EDTA, pH 8.0
Store ≤6 months at 4°C

M9 medium, 10×

30 g Na_2PO_4
15 g KH_2PO_4
5 g NH_4Cl
2.5 g NaCl
15 mg $CaCl_2$ (optional)
H_2O to 1 liter

Heat with stirring until dissolved. Pour aliquots into separate bottles, cap loosely, and autoclave 15 min at 15 lb/in^2. After the bottles cool to <50°C, add nutritional supplements and/or antibiotics as desired. After the bottles cool to <40°C, tighten bottle caps and store indefinitely at room temperature (discard if medium becomes contaminated).

M9ZB medium

Dissolve 10 g N-Z-Amine A (Sigma) and 5 g NaCl in 889 ml water. Autoclave, cool, and add 100 ml of 10× M9 medium (see recipe), 1 ml of 1 M sterile $MgSO_4$, and 10 ml of 40% (w/v) glucose (filter sterilized). Store ≤1 year at room temperature.

BOSTON UNIVERSITY CENTRAL LIBRARY

MCAC buffers (store ≤6 months at 4°C)

MCAC-0 buffer:
20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
1 mM PMSF (phenylmethylsulfonyl fluoride)

MCAC-1000 buffer:
20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
1 M imidazole
1 mM PMSF

Add PMSF immediately before use from a 0.2 M stock in 100% ethanol stored at room temperature. All buffers used with Ni^{2+}-NTA resin contain high salt concentrations to reduce nonspecific electrostatic interactions between proteins and resin. Lower salt concentrations can be used but may lead to nonspecific binding of unwanted proteins to the resin.

MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, and MCAC-200 buffers:

These buffers (containing different concentrations of imidazole) are made by mixing MCAC-0 buffer and MCAC-1000 buffer in the appropriate ratios: e.g., for MCAC-60 buffer, 94:6 (v/v) MCAC-0/MCAC-1000.

MCAC-EDTA buffer

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
0.1 M EDTA, pH 8.0
1 mM PMSF
Store ≤6 months at 4°C

Protease inhibitor cocktail, 150×

Stock solutions:
2 mg/ml aprotinin in H_2O
1 mg/ml leupeptin in H_2O
1 mg/ml pepstatin A in methanol

Cocktail: Combine 1.5 vol of each stock solution with 5.5 vol sterile water.

Protease inhibitors can be obtained from Sigma or U.S. Biochemical. The stock solutions and cocktail are stable ≥6 months at –20°C.

COMMENTARY

Background Information

Metal-chelate affinity chromatography (also called immobilized metal affinity chromatography or IMAC) for protein purification was first described in 1975 (Porath et al., 1975). It is based on the ability of certain amino acids acting as electron donors on the surface of proteins (histidine, tryptophan, tyrosine, or phenylalanine) to bind reversibly to transition-metal ions that have been immobilized by a chelating group covalently bound to a solid support. Of these amino acids, histidine is quantitatively the most important in mediating the binding of most proteins to immobilized metal ions. Histidine binds selectively to immobilized metal ions even in the presence of excess free metal ions in solution (Hutchens and Yip, 1990b). Copper and nickel ions have the greatest affinity for histidine (Yip et al., 1989).

Proteins and peptides containing phosphoamino acids (phophoserine, phosphotyrosine,

BCU IAS/CENTRAL UNIVERSITY LIBRARY

or phosphothreonine) have also been purified by MCAC via selective interaction of the terminal phosphate group with Fe^{3+} (Muszynska et al., 1992). Phosphoproteins also bind to less common metals such as lutetium, scandium, and thorium (Andersson and Porath, 1986).

The chelating group that has been used most extensively for MCAC is iminodiacetic acid (IDA). The tridentate IDA group binds to three sites within the coordination sphere of divalent metal ions such as copper, nickel, zinc, and cobalt. When copper ions (coordination number of 4) are bound to IDA, only one site remains available for interaction with proteins (Hochuli et al., 1987). For nickel ions (coordination number of 6) bound to IDA, three sites are available for binding to proteins. Thus Cu^{2+}-IDA complexes are stable on the column but have lower capacity for protein binding. Conversely, Ni^{2+}-IDA complexes bind proteins more avidly, but the Ni^{2+}-protein complexes are more likely to dissociate from the solid support.

A pentadentate chelating group—*N*,*N*,*N*′-tris(carboxymethyl)ethylenediamine (TED)—has also been used for MCAC. This chelator binds five coordination sites of the metal ions, providing highly stable metal ion–TED groups. However, for metal ions with a coordination number of 6, this leaves only one site available for protein interaction.

The development of a new metal-chelating adsorbent, nitrilotriacetic acid (NTA), has provided a convenient and inexpensive tool for purification of proteins containing histidine residues (Hochuli et al., 1987). The quadridentate NTA moiety covalently coupled via a spacer arm to Sepharose CL-6B forms an extremely stable complex with metal ions possessing a coordination number of 6 (e.g., Ni^{2+}). After binding to the NTA, the Ni^{2+} ion has two sites within its octahedral coordination sphere available for binding to electron donor groups (i.e., histidine) on the surface of proteins. Thus the advantage of NTA over IDA is that the Ni^{2+} ion is bound by four rather than three of its coordination sites. This minimizes leaching of the metal from the solid support and allows for more stringent purification conditions. The NTA also binds Cu^{2+} ions with high affinity, but this occupies all of the coordination sites, rendering the resulting complex ineffective for MCAC.

The affinity of histidine residues for immobilized Ni^{2+} ions allows selective purification of proteins containing a high proportion of histidine residues on the surface. Immobilized copper or nickel ions bind native proteins with a K_d of $1–17 \times 10^{-5}$ M (Hutchens and Yip, 1990a). This affinity is enhanced significantly by designing proteins to contain short stretches (6 to 10 residues) of histidines in regions likely to be surface exposed (amino or carboxyl termini) using recombinant DNA technology. Addition of a histidine tail results in a protein that binds to the Ni^{2+}-NTA complex with a K_d of 10^{-13} M at pH 8.0 even in the presence of detergent, ethanol, 2 M KCl (Hoffmann and Roeder, 1991), 6 M guanidine (Hochuli et al., 1988), or 8 M urea (Stüber et al., 1990).

The histidine tail binds to the Ni^{2+}-NTA resin via the imidazole side chains of the histidine residues. At pH ≥7.0, the imidazole side chain is deprotonated, with a net negative charge; at pH 5.97 (the pK of the imidazole side chain of histidine), 50% of the histidines are protonated; and at pH ≤4.5, almost all of the histidines are protonated and do not interact with Ni^{2+}-NTA. Thus, there are two methods of dissociating histidine tails from the Ni^{2+}-NTA resin. The first method, presented in this unit, uses increasing concentrations of imidazole at constant pH to displace the histidine tail from the Ni^{2+}-NTA. This technique provides great versatility because it can be used for both native and denaturing MCAC and buffer preparation is simplified. The second method uses buffers of decreasing pH to elute the histidine tail and is quite efficient for producing pure protein (Hochuli et al., 1988). Disadvantages are that the pH must be maintained accurately at all temperatures and that some proteins may not be able to withstand the extreme pH change.

Creation of a small histidine tail has advantages over other fusion protein systems. Addition of six histidines to the protein adds only 0.84 kDa to the mass of the protein, whereas other fusion protein systems utilize much larger affinity groups that often must be removed to allow normal protein function (e.g., glutathione-*S*-transferase, protein A, maltose-binding protein, and lacZ have masses of 26, 30, 40, and 116 kDa, respectively). The small histidine tail is not immunogenic and therefore need not be removed before the purified protein is injected into animals for antibody production. Histidine tail fusion proteins often retain their normal biologic functions—e.g., dihydrofolate reductase and adenylylcyclase retain their enzymatic activities (Hochuli, 1990; Taussig et al., 1993) and the TATA box binding protein retains its transcriptional regulatory function (Parvin et al., 1992). The aforementioned stud-

BOARD OF TRUSTEES LIBRARY

ies also demonstrate the technique's suitability for purification of membrane, cytoplasmic, and nuclear proteins. However, although it often is not necessary to remove the histidine tail, it is possible to include this step and still generate sufficient quantities of protein pure enough for detailed X-ray crystallographic studies (Nikolov et al., 1992).

The histidine tail can also be used to detect the recombinant protein. Ni^{2+}-NTA covalently coupled to alkaline phosphate (available from Qiagen; see *SUPPLIERS APPENDIX*) can bind directly to histidine tailed proteins, and the bound alkaline phosphatase can be quantitated by standard techniques (*UNIT 10.10*). This allows for direct quantitation of histidine-tailed proteins blotted to membranes without the need for primary and secondary antibodies (Botting and Randall, 1995).

The protocols presented in this unit utilize protein produced in bacteria; an example protocol for the purification of HIV-1 integrase expressed in *E. coli* is presented in *UNIT 6.5*. Histidine-tail fusion proteins have also been expressed and purified from HeLa cells using transient transfection techniques (Janknecht and Nordheim, 1992; Papavassiliou et al., 1992) or a vaccinia expression system (Janknecht et al., 1991) and from baculovirus-infected insect cells (Gearing et al., 1993; Taussig et al., 1993). Therefore, MCAC can be used with any protein expression system currently available.

This technique can also be used to prepare immunoaffinity columns (*UNIT 9.5*). Histidine tails have been covalently linked to oligosaccharide moieties of monoclonal antibodies and then subsequently bound to Ni^{2+}-NTA columns (Loetscher et al., 1992). The advantage of this method over more conventional techniques of antibody immobilization is that the immobilized antibody will retain a much greater ability to bind antigen and is recoverable from the column.

Critical Parameters

NTA can be synthesized according to the protocol provided by Hochuli (1990); more commonly, it is purchased from Qiagen, which provides detailed protocols for its use.

Binding and elution of proteins from the Ni^{2+}-NTA resin depend upon the histidine content of the protein. Each protein will have a slightly different histidine content and optimal elution profile on the Ni^{2+}-NTA resin. Therefore, it is important to establish the buffer imidazole concentrations that will give the best purification of a particular protein. As outlined in the Basic Protocol, gradient or stepwise elution of small-scale preparations should be done before large-scale preparations are attempted. It is also possible to perform batch purification by adding charged resin to a tube containing the extract, mixing for 1 hr, loading into a column, washing, and eluting.

Troubleshooting

The most common problem is low yield of purified protein in the expected fraction. This can be caused by insufficient protein loaded on the column, protein not binding to the column, or protein not eluting from the column.

Insufficient protein

It is important to verify that the desired protein is adequately expressed in the chosen system. Check the crude extract for the presence of the expressed protein by SDS-PAGE or other methods before proceeding to the affinity purification steps.

Protein in the crude lysate may undergo significant proteolytic degradation in some cell types. Always use PMSF, minimize freezing and thawing, maintain low temperatures during purification, and use additional protease inhibitors whenever practical.

Lack of binding

If the protein does not bind to the column, either the protein lacks a histidine tail, the column is inadequately charged with nickel, or the histidine tail is buried in the protein and is inaccessible to the resin. Before expression and purification are attempted, verify that the DNA sequence of the histidine tail and adjacent protein-coding sequence are in frame with each other and that initiation and termination codons are present. If the histidine tail is at the amino terminus, it may be helpful to delete the endogenous codon for initiator methionine; if this is present, translation may initiate downstream of it, producing a full-length protein without a tail. Placing the histidine tail at the carboxyl terminus ensures that only the full-length protein will bind to the Ni^{2+}-NTA resin (if protease activity is controlled). However, a histidine tail at the carboxyl terminus is more likely to be inaccessible to the resin.

Chelating agents (e.g., EDTA and EGTA) must be excluded from all solutions because they will strip the Ni^{2+} ions from the affinity column. Strong reducing agents such as dithiothreitol should also be avoided because they reduce the Ni^{2+} ions to metallic nickel,

BCU IAS / CENTRAL UNIVERSITY LIBRARY

forming a brown precipitate. Some proteins will require the presence of reducing agents to maintain structure. It is possible to use up to 10 mM 2-mercaptoethanol with the Ni^{2+}-NTA resin, although the lowest possible concentration should be used.

The Ni^{2+}-NTA should retain a distinct blue-green color throughout the purification process until EDTA is applied. If the resin is white or if the blue-green color is faint after charging the column with Ni^{2+}, check all buffers and make sure no chelating and reducing agents are present. Recharge the column with Ni^{2+} and if the color is still faint or absent, regenerate the NTA resin (see Support Protocol 2) and repeat Ni^{2+} charging (see Basic Protocol, steps 6 to 8). If the resin is still white, discard it and start over with a new batch.

If the protein has or should have a histidine tail but is not binding to the resin, perform the purification under denaturing conditions (see Alternate Protocol 1 or 2) to unmask the inaccessible histidine tail. Placing the histidine tail at the opposite terminus of the protein may also be helpful.

Failure to elute

If the protein appears to be binding to the column but not eluting, make a final elution with EDTA-containing buffer. Any protein bound to the Ni^{2+}-NTA will be removed with the EDTA. Occasionally the protein bound to the resin will precipitate and not elute; if this problem is suspected, purify the protein under denaturing conditions and analyze all eluted fractions by SDS-PAGE for the presence of the expressed protein.

Anticipated Results

The maximum binding capacity of Ni^{2+}-NTA resin is 5 to 10 mg of protein per milliliter of packed resin. Under ideal conditions, as much as 80% of the bound protein can be recovered. The limiting factor in most instances will probably be the amount of protein loaded on the resin, which depends on the protein expression system.

Time Considerations

Expression of protein and preparation of crude extract require 1 day for the system described here. Column preparation requires 30 to 60 min. Loading, washing, and eluting the column require 4 to 6 hr.

Literature Cited

Andersson, L. and Porath, J. 1986. Isolation of phosphoproteins by immobilized metal (Fe^{3+}) affinity chromatography. *Anal. Biochem.* 154:250-254.

Botting, C.H. and Randall, R.E. 1995. Reporter enzyme–nitrilotriacetic acid–nickel conjugates: Reagents for detecting histidine-tagged proteins. *Biotechniques*19:362-363.

Gearing, K., Göttlicher, M., Teboul, M., Widmark, E., and Gustafsson, J. 1993. Interaction of the peroxisome-proliferator-activated receptor and retinoid X receptor. *Proc. Natl. Acad. Sci. U.S.A.* 90:1440-1444.

Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R., and Stüber, D. 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* 6:1321-1325.

Hochuli, E., Dobeli, H., and Schacher, A. 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* 411:177-184.

Hoffmann, A. and Roeder, R. 1991. Purification of His-tagged proteins in nondenaturing conditions suggests a convenient method for protein interaction studies. *Nucl. Acids Res.* 19:6337-6338.

Hutchens, T.W. and Yip, T.-T. 1990a. Protein interactions with immobilized transition metal ions: Quantitative evaluations of variations in affinity and binding capacity. *Anal. Biochem.* 191:160-168.

Hutchens, T.W. and Yip, T.-T. 1990b. Differential interaction of peptides and protein surface structures with free metal ions and surface-immobilized metal ions. *J. Chromatogr.* 500:531-542.

Janknecht, R., De Martynoff, G., Lou, J., Hipskind, R., Nordheim, A., and Stunnenberg, H. 1991. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci. U.S.A.* 88:8972-8976.

Janknecht, R. and Nordheim, A. 1992. Affinity purification of histidine-tagged proteins transiently produced in HeLa cells. *Gene* 121:321-324.

Loetscher, P., Mottlau, L., and Hochuli, E. 1992. Immobilization of monoclonal antibodies for affinity chromatography using a chelating peptide. *J. Chromatogr.* 595:113-119.

Muszynska, G., Dobrowolska, G., Medin, A., Ekman, P., and Porath, J.O. 1992. Model studies on iron(III) affinity chromatography. II. Interaction of immobilized iron(III) ions with phosphorylated amino acids, peptides and proteins. *J. Chromatogr.* 604:19-28.

Nikolov, D., Hu, S., Lin, J., Gasch, A., Hoffman, A., Horikoshi, M., Chua, N., Roeder, R., and Burley, S. 1992. Crystal structure of TFIID TATA-box binding protein. *Nature* 360:40-46.

BOSTON UNIVERSITY LIBRARY

Papavassiliou, A., Bohmann, K., and Bohmann, D. 1992. Determining the effect of inducible protein phosphorylation on the DNA-binding activity of transcription factors. *Anal. Biochem.* 203:302-309.

Parvin, J., Timmers, H., and Sharp, P. 1992. Promoter specificity of basal transcription factors. *Cell* 68:1135-1144.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258:598-599.

Struhl, K. 1987. Subcloning of DNA fragments. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 3.16.1-3.16.11. John Wiley & Sons, New York.

Stüber, D., Matile, H., and Garotta, G. 1990. System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: Application to epitope mapping, preparation of antibodies, and structure-function analysis. *Immunol. Methods* 4:121-152.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 polymerase to direct expression of cloned genes. *Methods Enzymol.* 185:60-89.

Taussig, R., Quarmby, L., and Gilman, A. 1993. Regulation of purified type I and type II adenylylcyclases by G protein β-γ subunits. *J. Biol. Chem.* 268:9-12.

Yip, T.-T., Nakagawa, Y., and Porath, J. 1989. Evaluation of the interaction of peptides with Cu(II), Ni(II), and Zn(II) by high-performance immobilized metal ion affinity chromatography. *Anal. Biochem.* 183:159-171.

Key Reference

Hochuli, E. 1990. Purification of recombinant proteins with metal chelate adsorbent. *In* Genetic Engineering, Principles and Practice, Vol. 12 (J. Setlow, ed.) pp. 87-98. Plenum, New York.

Describes basic principles of MCAC with detailed protocols.

Contributed by Kevin J. Petty
University of Texas Southwestern Medical Center
Dallas, Texas

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Immunoaffinity Chromatography

This unit describes the isolation of soluble or membrane-bound protein antigens from cells or homogenized tissue by immunoaffinity chromatography. This technique involves the elution of a single protein from an immunoaffinity column after prior elution of nonspecifically adsorbed proteins. Specifically, antibodies are coupled to Sepharose (an insoluble, large-pore-size chromatographic matrix). High-molecular-weight antigens pass freely into and out of the pores and bind to antibodies covalently bound to the matrix. To elute the bound antigen from the immunoaffinity matrix, the antibody-antigen interaction is destabilized by brief exposure to high-pH (Basic Protocol) or low-pH (Alternate Protocol 1) buffer. The use of batch purification of antigens shortens the column loading time (Alternate Protocol 2). The detergent octyl β-D-glucoside can be used instead of Triton X-100 for elution. Because octyl β-D-glucoside has a high critical micelle concentration (CMC) it can be readily removed by dialysis (Alternate Protocol 3). The procedure for covalently linking an antibody to Sepharose using the cyanogen bromide activation method is given in the Support Protocol.

ISOLATION OF SOLUBLE OR MEMBRANE-BOUND ANTIGENS

BASIC PROTOCOL

Two different Sepharose columns in series—a precolumn to remove nonspecifically binding material and a specific column—are used to isolate antigens from a cell or tissue lysate. Column fractions are analyzed by SDS-PAGE and silver staining to detect the antigens.

Materials

Antibody (Ab)-Sepharose (see Support Protocol)
Activated, quenched (control) Sepharose, prepared as for Ab-Sepharose (see Support Protocol) but eliminating Ab or substituting irrelevant Ab during coupling
Cells or homogenized tissue
Tris/saline/azide (TSA) solution (see recipe), ice cold
Lysis buffer (see recipe), ice cold
5% (w/v) sodium deoxycholate (Na-DOC; filter sterilize and store at room temperature)
Wash buffer (see recipe)
Tris/Triton/NaCl buffers, pH 8.0 and 9.0 (see recipe), ice-cold
Triethanolamine solution (see recipe), ice cold
1 M Tris·Cl, pH 6.7 (*APPENDIX 2E*), ice cold
Column storage solution (see recipe), ice cold

Chromatography columns
Ultracentrifuge
Quick-seal centrifuge tubes (Beckman)

Additional reagents and equipment for column chromatography (*UNITS 8.1*), SDS-PAGE (*UNIT 10.1*), and silver staining (*UNIT 10.5*)

NOTE: Carry out all procedures involving antigen in a 4°C cold room or on ice.

Prepare the columns

1. Prepare an Ab-Sepharose immunoaffinity column (5 ml; 5 mg/ml antibody per milliliter packed Sepharose) and an activated, quenched (control) Sepharose precolumn (5 ml packed bed volume) linked in series (Fig. 9.5.1).

Irrelevant antibody can be coupled to the Sepharose in the precolumn.

Column size can vary; adjust amounts of Sepharose and cells proportionally.

Copyright © 1996 by John Wiley & Sons, Inc.

BOWLING GREEN STATE UNIVERSITY LIBRARY

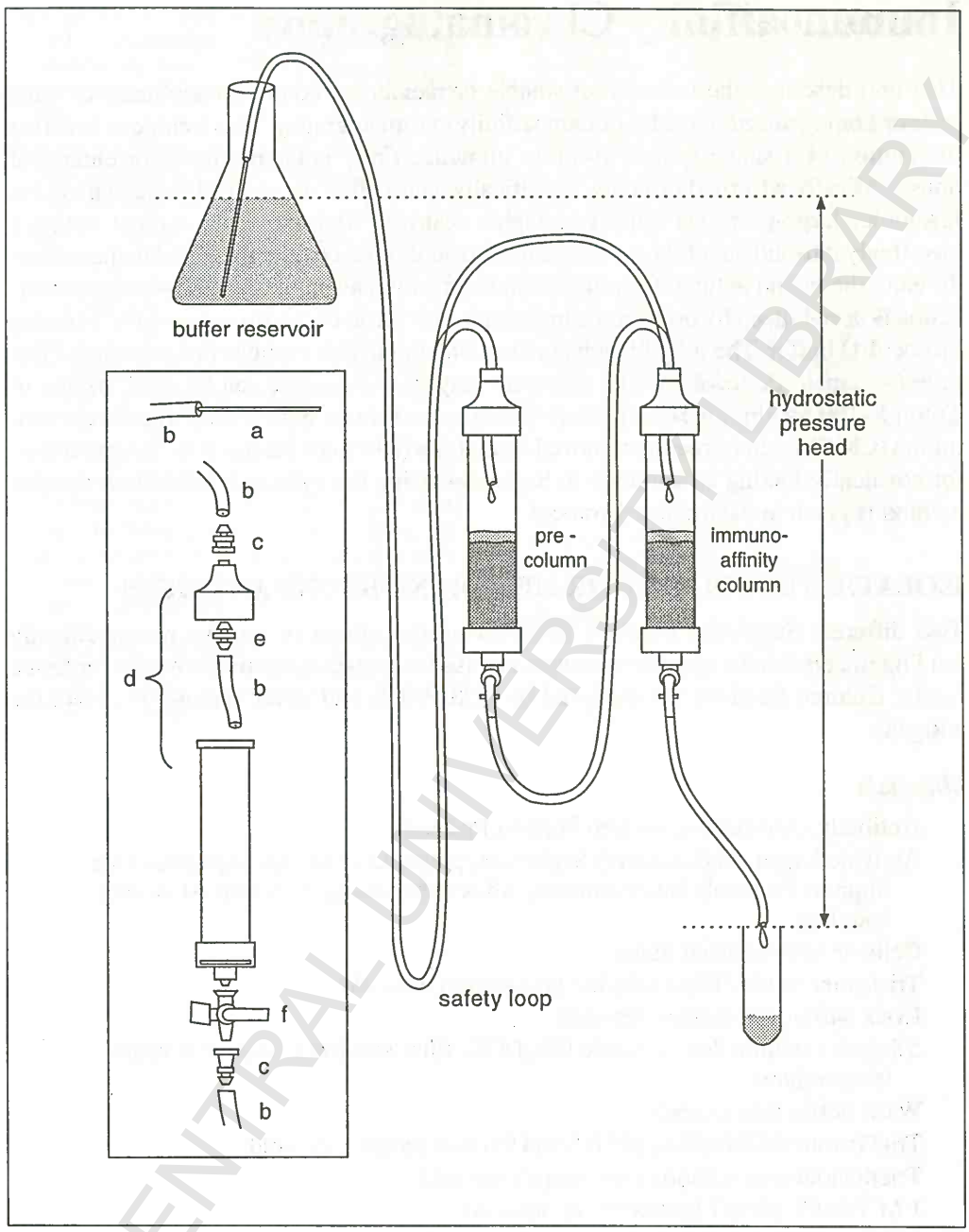


Figure 9.5.1 Immunoaffinity chromatography. During the application of the sample, two Sepharose columns, a Sepharose precolumn (without covalently bound specific antibody or with a covalently bound irrelevant antibody) and an immunoaffinity column (with covalently bound antibody), are attached in series to a buffer reservoir containing the sample. After the sample has been washed through, the precolumn is removed and the tubing of the safety loop is connected to the immunoaffinity column. The hydrostatic pressure head is the distance between the top of the solution in the buffer reservoir and the tip of the tubing at the bottom of the immunoaffinity column. When the elution reservoir is emptied, the hydrostatic head becomes zero when the fluid level reaches the safety loop, preventing columns from running dry. Fluid remaining above the column beds can be removed by raising the safety loop. After rinsing the tubing, the next elution is begun by placing the end of the safety loop in another reservoir containing the next elution buffer.

Inset: Schematic diagram of an immunoaffinity column. (a) 50-µl disposable capillary micropipet. (b) Tubing: Tygon S-54-HL Microbore, 0.05-in. i.d., or Tygon R-3603, 1/16-in. i.d. (softer tubing). (c) Female Luer fitting, white nylon (Value Plastics), 1/16 in. (d) Kontes Flex-column (Kontes Glass). (e) Barbed nipple connector, polypropylene, 3/32-in. top, 1/16-in. bottom (Value Plastics, Series AD). (f) Luer-Lok two-way stopcock (Kontes Glass).

BCU IAS CENTRAL UNIVERSITY LIBRARY

Prepare the lysate

2. Suspend 50 g of cells at 1–5 × 10^8 cells/ml in ice-cold TSA solution, or add 1 to 5 vol ice-cold TSA per vol packed cells or homogenized tissue. Add an equal volume of ice-cold lysis buffer and stir 1 hr at 4°C.

   For glycosylphosphatidylinositol (GPI)-anchored proteins (UNIT 12.5), incubate 10 min at 20°C for efficient solubilization.

3. Centrifuge 10 min at 4000 × *g*, 4°C, to remove nuclei. Decant supernatant and save.

   For purification of cytoplasmic (soluble) antigens, it is not necessary to add detergents to the solutions and buffer used in subsequent steps. Detergent is needed only for cell lysis and solubilization of integral membrane proteins.

4. For purification of membrane antigens, add 0.2 vol of 5% Na-DOC to the postnuclear supernatant, and leave 10 min at 4°C or on ice. Transfer to quick-seal centrifuge tubes and centrifuge 1 hr at 100,000 × *g*, 4°C. Carefully remove supernatant and save.

Set up and wash the columns

5. Attach Sepharose precolumn to immunoaffinity column (Fig. 9.5.1).

6. Wash both columns using the following regimen:

   10 column volumes of wash buffer
   5 column volumes of Tris/Triton/NaCl buffer, pH 8.0
   5 column volumes of Tris/Triton/NaCl buffer, pH 9.0
   5 column volumes of triethanolamine solution
   5 column volumes of wash buffer.

Isolate the antigen

7. Apply the supernatant from steps 3 or 4 (reserving some for analysis as described in step 15 below) to the precolumn and allow it to flow through the precolumn and specific column linked in series at a flow rate of 5 column volumes/hr. Collect the flow-through fractions, each 1/10 to 1/100 the volume of the applied supernatant.

   "Fat" chromatography columns, filled with Sepharose to a height of ~2× column diameter, are used to maximize flow rates. A 10- to 20-ml syringe is used for 5 ml of Sepharose. The flow rate is adjusted with a hydrostatic head of up to 250 cm (Fig. 9.5.1). Sample loading can routinely take up to 2 days with no deleterious effect, but longer periods would suggest the column is clogged or the lysate is too viscous. The latter is usually due to the presence of DNA.

8. Wash with 5 column volumes of wash buffer, then close the stopcocks on both columns and disconnect the precolumn from the immunoaffinity column. Open the stopcock of the immunoaffinity column and allow fluid above the top of the column to drain out to bed level.

   The Sepharose has some elasticity and draining can continue until there is no buffer above the Sepharose bed. Draining until cracks appear in the Sepharose should be avoided.

   Fractions of this wash and washes obtained below should be saved.

9. Wash the immunoaffinity column between each change of buffers (steps 10 to 14) as follows. Close the stopcock and remove the end cap of the column. With a syringe connected to the outlet of the tubing from the buffer reservoir, aspirate all buffer from the tubing. Place tubing into the next buffer contained in another reservoir. Aspirating with a syringe, fill the tubing from the reservoir and remove the syringe. Crimp the tubing to regulate flow and rinse the inside wall of the column with the buffer. Open

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

the column stopcock and drain the buffer to bed level. Put end cap loosely on the column and allow buffer to drain into the column to a level several centimeters above the bed. Secure end cap and commence washes or elution.

10. Wash with 5 column volumes of wash buffer.
11. Wash with 5 column volumes of Tris/Triton/NaCl buffer, pH 8.0.
12. Wash with 5 column volumes of Tris/Triton/NaCl buffer, pH 9.0.

    Some nonspecifically bound proteins may be eluted at this step. In addition, some monoclonal antibodies may release some ligand at this pH. This should be checked.

13. Elute the antigen with 5 column volumes of triethanolamine solution. Collect fractions of 1 column volume into tubes containing 0.2 vol of 1 M Tris·Cl, pH 6.7, to neutralize the fractions collected.

    In some cases it may be desirable to lower the pH of the triethanolamine solution to preserve the functional activity of the ligand. The ideal pH gives complete release of the ligand, as verified by SDS-PAGE evaluation of a sample (~20 µl) of the eluted column bed (Ab-Sepharose) and eluate (50 µl).

14. Wash the column with 5 column volumes of TSA solution.

    A column may be reused many times and remain active for several years after storage at 4°C in TSA solution. It is important to prevent drying out of a column during storage. The use of column storage solutions inhibits the growth of microorganisms.

15. Analyze fractions for the presence of antigen—50-µl aliquots of each eluate fraction should be analyzed by SDS-PAGE and silver staining. Analyze 0.5- to 1-ml aliquots of the sample applied to the column and representative flowthrough and wash fractions by immunoprecipitation with Ab-Sepharose, and detect by silver staining to determine whether the column was saturated.

    If antibody leaches off the column during elution, it may be removed from the eluate by passage through protein A–Sepharose (Ey et al., 1978). Even the weakly binding mouse IgG1 subclass can be quantitatively removed at pH 8 (M. Dustin, pers. comm.).

    Preliminary analysis of fractions can be done using A_{280} readings when octyl β-D-glucoside or sodium deoxycholate are used as detergents.

ALTERNATE PROTOCOL 1

LOW-pH ELUTION OF ANTIGENS

Some protein antigens may be eluted more completely with greater retention of native conformation and with fewer contaminants when low-pH buffers are employed.

Additional Materials *(also see Basic Protocol)*

Sodium phosphate buffer, pH 6.3 (see recipe)
Glycine buffer (see recipe)
1 M Tris·Cl, pH 9.0 (*APPENDIX 2E*)

1. Prepare the columns and lysate, wash the columns, and isolate the antigen (see Basic Protocol, steps 1 to 11).

   It is essential to remove sodium deoxycholate from the column before acid elution, because it precipitates or forms a gel at acid to neutral pH.

   TSA and triethanolamine solutions used here may also be modified by substituting 1% octyl β-D-glucoside for 0.1% Triton X-100 (see Alternate Protocol 3).

2. Wash with 5 column volumes of sodium phosphate buffer, pH 6.3.

BCU IASI CENTRAL UNIVERSITY LIBRARY

3. Elute with 5 column volumes of glycine buffer. Collect fractions into tubes containing 0.2 vol of 1 M Tris·Cl, pH 9.0.

   Mix each fraction immediately after collection.

4. Analyze fractions for antigen (see Basic Protocol, step 15).

BATCH PURIFICATION OF ANTIGENS

ALTERNATE PROTOCOL 2

Batch purification of antigens shortens the column loading time. This technique is valuable for viscous lysates that take too long to load on a column and for antigens especially susceptible to proteolysis, because less time is required to complete the steps. A precolumn is not utilized because the supernatant is mixed with Ab-Sepharose and poured into a column. The antigen is then eluted (see Basic Protocol). The drawbacks of this protocol are that more "hands-on" time is required by the investigator and that nonspecifically binding material is not removed by a precolumn.

1. Suspend and centrifuge the cells and purify the membrane antigens (see Basic Protocol, steps 2 to 4) to obtain the postnuclear supernatant.
2. Suspend Ab-Sepharose in the supernatant in a flask. Shake gently on a rotary shaker for 3 hr. Stop shaking and allow the Sepharose to settle.
3. Decant most of the supernatant. Pour the Ab-Sepharose and the remainder of the supernatant into a column and open the stopcock. Continue draining the column until all the Sepharose has been added. Allow the fluid to drain to bed level and close the stopcock.
4. Wash the immunoaffinity column, elute the antigen, and analyze the fractions (see Basic Protocol, steps 9 to 15).

ELUTION IN OCTYL β-D-GLUCOSIDE

ALTERNATE PROTOCOL 3

The detergent octyl β-D-glucoside has a high critical micelle concentration (CMC) of 0.73% and can therefore be readily removed by dialysis. Also, adsorption of membrane proteins to surfaces for ELISA or adhesion assays is much more efficient when solutions containing membrane proteins are diluted so that the concentration of octyl β-D-glucoside is below the CMC. Because octyl β-D-glucoside is expensive, initial steps requiring large volumes may be done as described in the Basic Protocol; the initial detergent is then replaced by octyl β-D-glucoside in the wash step prior to elution.

Additional Materials (also see Basic Protocol)

TSA solution (see recipe) containing 1% octyl β-D-glucoside

1. Prepare the columns and lysate, wash the columns, and isolate the antigen (see Basic Protocol, steps 1 to 11).
2. Wash with 5 column volumes of TSA solution containing 1% octyl β-D-glucoside.
3. Elute with 5 column volumes of triethanolamine solution with 1% octyl β-D-glucoside substituted for 0.1% Triton X-100.
4. Collect fractions of 1 column volume into tubes containing 0.2 vol of 1 M Tris·Cl, pH 6.7, to neutralize the fractions collected.
5. Wash the column and analyze the fractions (see Basic Protocol, steps 14 and 15).

BCU IFSI CENTRAL UNIVERSITY LIBRARY

PREPARATION OF ANTIBODY-SEPHAROSE

This protocol details the procedure for covalently linking an antibody to Sepharose (an insoluble, large-pore-size chromatographic matrix) using the cyanogen bromide activation method (also see *UNIT 9.3*). First, the antibody and Sepharose are prepared separately. The Sepharose is then activated with cyanogen bromide (alternatively, CNBr-activated Sepharose can be purchased from Pharmacia Biotech and used according to the manufacturer's instructions). Finally, the CNBr-activated Sepharose is coupled to the antibody.

CAUTION: CNBr is very toxic, releases toxic vapor, and is reported to create explosive side products upon storage; therefore, it should always be used in a hood and handled with appropriate CNBr- and solvent-resistant gloves. Only a fresh bottle of CNBr should be employed and it should be opened in the hood. Because of the problems involved in handling CNBr, it is best to dissolve a complete bottle in a defined amount of anhydrous acetonitrile. Short-term storage of this solution (several weeks) at 0°C is possible, but prolonged storage is not recommended.

Materials

1 to 30 mg/ml antigen-specific monoclonal or polyclonal antibody
0.1 M $NaHCO_3$/0.5 M NaCl
Sepharose CL-4B (or Sepharose CL-2B for high-molecular-weight antigens; Pharmacia Biotech)
2 M Na_2CO_3
Cyanogen bromide (CNBr)/acetonitrile (see recipe)
1 mM and 0.1 mM HCl, ice-cold
0.05 M glycine (*or* ethanolamine), pH 8.0
Tris/saline/azide (TSA) solution (see recipe)

Dialysis tubing (MWCO >10,000)
Ultracentrifuge
Whatman no. 1 filter paper
Buchner funnel
Erlenmeyer filtration flask
Water aspirator

Prepare the antibody

1. Dialyze 1 to 30 mg/ml antibody against 0.1 M $NaHCO_3$/0.5 M NaCl at 4°C with three buffer changes during 24 hr. Use a volume of dialysis solution that is 500 times the volume of antibody solution.

   Dialysis is performed to remove all small molecules containing free amino or sulfhydryl groups.

2. Centrifuge 1 hr at 100,000 × *g*, 4°C, to remove aggregates. Save the supernatant.

   Removal of aggregates is important. Because only some of the antibody molecules in an aggregate will directly couple to the Sepharose, the noncoupled antibody molecules may leach out during elution.

3. Measure the A_{280} of an aliquot of the solution and determine the concentration of the antibody (mg/ml IgG = A_{280}/1.44). Dilute with 0.1 M $NaHCO_3$/0.5 M NaCl to 5 mg/ml (or to the same concentration as desired for Ab-Sepharose) and keep at 4°C. Measure the A_{280} of this solution for comparison of absorbance in step 10.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Prepare the Sepharose

4. Allow the Sepharose slurry to settle in a beaker or the container in which it is sold; decant and discard the supernatant. Weigh out the desired quantity of Sepharose (assume density = 1.0).

5. Set up a filter apparatus using Whatman no. 1 filter paper in a Buchner funnel and an Erlenmeyer filtration flask attached to a water aspirator. Wash the Sepharose on the filter apparatus with 10 vol water.

   Sintered-glass funnels are traditionally recommended but rapidly become clogged unless coarse-porosity funnels are used.

Activate Sepharose with CNBr

6. Transfer Sepharose to 50-ml beaker and add an equal volume of ice-cold 2 M Na_2CO_3.

7. Activate Sepharose in an ice bath using 3.2 ml CNBr/acetonitrile per 100 ml Sepharose. Add CNBr/acetonitrile dropwise with a Pasteur pipet over 1 min, while slowly stirring the slurry with a magnetic stir bar. Continue stirring slowly for 5 min.

   Excessive and vigorous stirring may fracture the Sepharose beads; this may cause flow problems during column chromatography. The protocol uses 2 g CNBr/100 ml Sepharose. Two to four grams of CNBr/100 ml Sepharose can be used to couple 1 to 20 mg of antibody/ml Sepharose.

   CAUTION: *Activation should be carried out in a fume hood.*

8. Rapidly filter the CNBr-activated Sepharose as in step 5. Aspirate to semidryness (i.e., until the Sepharose cake cracks and loses its sheen). Wash with 10 vol ice-cold 1 mM HCl, then with 2 vol ice-cold 0.1 mM HCl. Hydrate with enough ice-cold 0.1 mM HCl so that the cake regains its sheen, but there is no excess liquid above the cake.

   Washing is most efficient if the wash solution is added evenly over the surface of the cake at about the same rate as the solution is removed by filtration. CNBr-activated Sepharose is very unstable at the alkaline pH necessary for activation; it is much more stable in dilute HCl. CNBr-activated Sepharose can be purchased from Pharmacia Biotech, but the coupling capacity will be lower.

   Destroy the contents of the filter flask, which contains excess CNBr, by slowly adding a 5-fold molar excess (relative to the total initial amount of CNBr used) of 1 M NaOH. Keep the reaction mixture cold by adding pieces of ice. Perform this decomposition in the hood, and keep the mixture there until all CNBr is destroyed.

Couple antibody to CNBr-activated Sepharose

9. Immediately transfer a weighed amount of Sepharose (assume density = 1.0) to a beaker. Add an equal volume of a solution of antibody dissolved in 0.1 M $NaHCO_3$/0.5 M NaCl (from step 3). Stir gently with a magnetic stir bar or rotate end-over-end 2 hr at room temperature or overnight at 4°C.

10. Add 0.05 M glycine (or ethanolamine), pH 8.0, to saturate the remaining reactive groups on the Sepharose and allow the slurry to settle. Remove an aliquot of the supernatant and centrifuge to remove any residual Sepharose. Measure the A_{280} and compare to the A_{280} of the antibody solution from step 3 to determine the percentage coupling.

11. Store the Ab-Sepharose in TSA solution.

BCU IRSIC GEN RAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

CNBr/acetonitrile

To 25 g of cyanogen bromide in the original bottle *(CNBr should be white, not yellow, crystals)* add 50 ml acetonitrile to make a 62.5% (w/v) solution. Store indefinitely at –20°C in a desiccator over silica. Allow to warm before opening.

CAUTION: *CNBr is a highly toxic lachrymator; handle in a fume hood.*

Column storage solution

Prepare in TSA solution (see recipe) either 1 mM EDTA/20 µg/ml gentamicin *or* 0.01% thimerosal (Aldrich).

Detergent stock solutions

Prepare 10% Triton X-100 *or* 5% sodium deoxycholate in water. Sterilize either solution by Millipore filtration. Both solutions remain stable for 5 years at room temperature; Triton X-100 solution should be stored in the dark to prevent photo-oxidation.

NP-40 can be used in place of Triton X-100.

Glycine buffer

50 mM glycine·HCl, pH 2.5
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.15 M NaCl

Lysis buffer

TSA solution (see recipe) containing:
2% Triton X-100 (see recipe for detergent stock solutions)
5 mM iodoacetamide
Aprotinin (0.2 trypsin inhibitor U/ml)
1 mM phenylmethylsulfonyl fluoride (add fresh from 100 mM stock solution prepared in absolute ethanol)

IMPORTANT NOTE: *Iodoacetamide is a protease inhibitor and prevents oxidation of free cysteines to disulfide-bonded cysteines. It should be omitted for enzymes that require cysteines for activity.*

NP-40 can be used in place of Triton X-100.

Sodium phosphate buffer, pH 6.3

50 mM sodium phosphate, pH 6.3
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.5 M NaCl

Triethanolamine solution

50 mM triethanolamine, pH ~11.5
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.15 M NaCl

Other organic solutions, such as diethylamine, may be used in place of triethanolamine. The pH of this solution should be determined for each antibody as elution conditions may vary (see Critical Parameters).

BCU IASI CENTRAL UNIVERSITY LIBRARY

Tris/saline/azide (TSA) solution

0.01 M Tris·Cl, pH 8.0 (at 4°C)
0.14 M NaCl
0.025% NaN_3

CAUTION: *Sodium azide (NaN_3) is poisonous; wear gloves and handle cautiously.*

Tris/Triton/NaCl buffer, pH 8.0 and 9.0

50 mM Tris·Cl, pH 8.0 or pH 9.0
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.5 M NaCl

Wash buffer

0.01 M Tris·Cl, pH 8.0 (at 4°C)
0.14 M NaCl
0.025% NaN_3
0.5% Triton X-100 (see recipe for detergent stock solutions)
0.5% sodium deoxycholate (see recipe for detergent stock solutions)

CAUTION: *Sodium azide (NaN_3) is poisonous; wear gloves and handle cautiously.*

COMMENTARY

Background Information

The review of affinity chromatography by Wilchek et al. (1984) discusses available methods for activation of solid supports, coupling of ligands, adsorption of proteins, and elution of protein from affinity columns. Table III of that review lists numerous examples of proteins that have been purified by immunoaffinity chromatography and the elution conditions for each purification.

Traditionally, purification of membrane proteins started with a membrane purification step, and in some cases this is still desirable. However, it is difficult to achieve more than a 5-fold purification of plasma membranes and yields are usually only 10% to 40%. Omission of membrane purification in this protocol (Williams and Barclay, 1986; Johnson et al., 1985) results in increased yield and decreased experiment time.

Purification to homogeneity or near-homogeneity can usually be achieved for protein antigens present in ≥10,000 molecules per eukaryotic cell. This protocol can be used for both membrane and intracellular antigens. However, for soluble antigens, immunoaffinity chromatography is completed without detergent.

Critical Parameters

Binding capacities of Ab-Sepharose columns (coupled at 5 mg monoclonal antibody/ml Sepharose) have been found to be 2% to 20% of the theoretical binding capacity. The lowest and highest binding capacity values were found for antigens of 150,000 and 18,000 M_r, respectively, suggesting that increased antigen size may constrain access to antibody in the pores of the affinity matrix. Because of its larger pore size, Sepharose CL-4B (a 4% cross-linked agarose) is far preferable to Sepharose 6B (a 6% agarose), which is the usual commercially available preactivated agarose. The cross-links in the CL series yield mechanically more robust beads and do not appreciably decrease activation with cyanogen bromide (CNBr).

The Support Protocol for coupling protein antigens to CNBr-activated Sepharose is a modification of the methods of Cuatrecasas (1970) and March et al. (1974). As originally described, the washing was done at alkaline pH. Because activated Sepharose is very unstable at this pH, it was originally recommended that washing, adding the protein ligand, and mixing be done in <90 sec (Cuatrecasas, 1970). However, with the use of an acid wash buffer (Gelb, 1973) as described in the Support Protocol, activated groups remain stable for ≥30 min. Wilchek et al. (1984) described an alternative activation procedure with CNBr triethanolamine in acetone/water that is highly efficient and minimizes side-reaction derivatization of the Sepharose. The 0.1 M $NaHCO_3$ buffer in which antibody is dissolved provides an optimal pH of ~8.4 after mixing with activated Sepharose. The small amount of CNBr recommended is sufficient to obtain a coupling yield of 80% to 95%. Higher amounts of CNBr may result in multipoint attachment of IgG molecules to the matrix, thereby inactivating

B.C.U. „M. Eminescu" IAŞI
CHIMIE

BIBLIOTECA CENTRALĂ UNIVERSITY LIBRARY

antigen binding sites and reducing accessibility to antigen. Coupling efficiencies on the order of 99% are usually associated with lower antigen binding capacity and are to be avoided; coupling efficiencies of 80% to 90% are ideal. Smaller amounts of CNBr per milliliter of Sepharose also result in lower derivatization of the Sepharose with charged groups. These charged groups can cause nonspecific binding.

The amount of CNBr used in commercially available activated Sepharose is not specified, but is probably high. Many investigators purchase CNBr-activated Sepharose; others activate their own. Commercially activated Sepharose is stabilized by drying. Not all Sepharose can withstand drying; e.g., Sepharose 2B (2% agarose beads) is less robust than Sepharose 4B (4% agarose) and is not available in a preactivated form. Activating your own Sepharose allows more control over coupling efficiency; enables use of CL (cross-linked) Sepharose, which is more robust and has faster flow rates; enables use of Sepharose CL-2B; and is more cost effective.

A binding capacity of 40% was reported for coupling at 2 to 3 mg antibody/ml Sepharose. Successful purification has been achieved using monoclonal antibodies with affinity constants ranging from 2×10^7 to 4×10^8 M^{-1}. The column should be saturated with antigen by allowing some of the antigen to flow through the column during loading. This will result in the highest antigen purity and a mass yield, and will diminish the relative level of antibody eluted along with the antigen when antibody is leaching off the column.

Sodium deoxycholate is used in the solubilization protocol (Johnson et al., 1985) because it dissociates proteins from the membrane more effectively than Triton X-100. However, because sodium deoxycholate releases DNA from nuclei, it must be added to the lysate after the nuclei are removed. Sodium deoxycholate forms a mixed micelle with Triton X-100. Although sodium deoxycholate can be substituted for Triton X-100 during high-pH elution, it gels at low pH and in high salt. Both sodium deoxycholate and nonionic detergents may dissociate subunits of protein complexes, which interact within the membrane. Addition of phospholipid, low concentrations of Triton X-100, and mild detergents (e.g., digitonin, octylglucoside, and CHAPS) preserves membrane protein complexes (Helenius et al., 1979; Rivnay et al., 1982; Tsuchiya and Saito, 1984).

Glycosylphosphatidylinositol (GPI)-anchored proteins (*UNIT 12.5*) are resistant to detergent solubilization at 4°C (Schroeder et al., 1994). Preparation of the lysate (Basic Protocol, step 2) should be conducted for 10 min at 20°C for these antigens.

Octyl β-D-glucoside is an expensive detergent that is readily removable by dialysis. Furthermore, adsorption of proteins to substrates is enhanced by dilution below a CMC of 0.73%. In Alternate Protocol 3, the initial detergent is replaced with octyl β-D-glucoside prior to elution of antigen from the affinity column.

Protein antigens eluted by acid or base can frequently be renatured by neutralization. However, some protein antigens are irreversibly denatured. The structure of certain antigens is preserved after acid, but not base, elution (Plunkett and Springer, 1986), whereas the structure of other antigens is preserved after base, but not acid, elution (Johnson et al., 1985). Some antigens are eluted at low pH, but not at high pH; for others the reverse is true. For each antibody-antigen combination, the optimal pH for elution of specific antigens as well as contaminants must be empirically defined. Antibody binding capacity is usually retained after repeated exposure to low and high pH elution buffers (see Alternate Protocol 1 and see Basic Protocol, respectively). The use of chaotropic agents (e.g., potassium thiocyanate and urea) for elution is a seldom-employed alternative (Johnson et al., 1985).

Troubleshooting

Immunoaffinity chromatography relies on the elution of a single protein from an immunoaffinity column after prior elution of all other nonspecifically adsorbed proteins. Thus, depending on the elution conditions used, the desired protein antigen may be contaminated with other proteins. It is very important to determine whether a contaminant is present if the protein is to be analyzed for amino acid composition or protein sequence. One-dimensional gel electrophoresis (*UNIT 10.1*) should be used to verify elution of contaminating proteins during washing of the immunoaffinity column, as well as the purity of the protein in the final eluate. If the protein is not pure, the wash steps must be optimized to ensure that other contaminating proteins are removed.

Anticipated Results

Antigen yield is typically 40% to 70% of starting material (Kürzinger and Springer, 1982; Johnson et al., 1985) and purification factors of 1,000- to 10,000-fold may be achieved (Kürzinger and Springer, 1982; Williams and Bar-

BCU IAS CENTRAL LIBRARY

clay, 1986; Plunkett and Springer, 1986). Further purification can usually be achieved by a second cycle of immunoaffinity chromatography. Monoclonal antibodies are most convenient, but affinity-purified polyclonal IgG can also be used.

Time Considerations

Pouring the column takes a few minutes and lysate preparation requires ~6 hr. Purification proceeds over 1 to 2 days depending on the flow rate of the immunoaffinity column. The majority of this time involves loading the sample on the column, which may be done from a reservoir and requires little hands-on time. The use of batch purification (Alternate Protocol 2) reduces the sample application time to 3 hr. Elution of an immunoaffinity column requires 5 to 6 hr.

Literature Cited

Cuatrecasas, P. 1970. Protein purification by affinity chromatography. *J. Biol. Chem.* 245:3059-3065.

Ey, P.L., Prowse, S.J., and Jenkin, C.R. 1978. Isolation of pure IgG_1, IgG_{2a} and IgG_{2b} immunoglobulins from mouse serum using protein A–Sepharose. *Immunochemistry* 15:429-436.

Gelb, W.G. 1973. Affinity chromatography: For separation of biological materials. *Am. Lab.* 81:61-67.

Helenius, A., McCaslin, D.R., Fries, E., and Tanford, C. 1979. Properties of detergents. *Methods Enzymol.* 56:734-749.

Johnson, P., Williams, A.F., and Woollett, G.R. 1985. Purification of membrane glycoproteins with monoclonal antibody affinity columns. *In* Hybridoma Technology in the Biosciences and Medicine (T.A. Springer, ed.) pp. 163-175. Plenum, New York.

Kürzinger, K. and Springer, T.A. 1982. Purification and structural characterization of LFA-1, a lymphocyte function-associated antigen, and Mac-1, a related macrophage differentiation antigen. *J. Biol. Chem.* 257:12412-12418.

March, S.C., Parikh, I., and Cuatrecasas, P. 1974. A simplified method for cyanogen bromide activation of agarose for affinity chromatography. *Anal. Biochem.* 60:149-152.

Plunkett, M.L. and Springer, T.A. 1986. Purification and characterization of the lymphocyte function-associated-2 (LFA-2) molecule. *J. Immunol.* 136:4181-4187.

Rivnay, B., Wank, S.A., Poy, G., and Metzger, H. 1982. Phospholipids stabilize the interaction between the alpha and beta subunits of the solubilized receptor for immunoglobulin E. *Biochemistry* 21:6922-6927.

Schroeder, R., London, E., and Brown, D. 1994. Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. *Proc. Natl. Acad. Sci. U.S.A.* 91:12130-12134.

Tsuchiya, T. and Saito, S. 1984. Use of *n*-octyl-β-D-thioglucoside, a new nonionic detergent, for solubilization and reconstitution of membrane proteins. *J. Biochem.* 96:1593-1597.

Wilchek, M., Miron, T., and Kohn, J. 1984. Affinity chromatography. *Methods Enzymol.* 104:3-55.

Williams, A.F. and Barclay, A.N. 1986. Glycoprotein antigens of the lymphocyte surface and their purification by antibody affinity chromatography. *In* Immunological Methods in Biomedical Sciences (D.M. Weir, L.A. Herzenberg, C.C. Blackwell, and L.A. Herzenberg, eds.) pp. 22.1-22.24. Blackwell Scientific, Oxford.

Key References

Harlow, E. and Lane, D. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Hjelmeland, J.M. and Chrambach, A. 1984. Solubilization of functional membrane proteins. *Methods Enzymol.* 104:305-318.

Johnson et al., 1985. See above.

Describes the critical parameters involved in immunoaffinity chromatography.

Wilchek et al., 1984. See above.

Describes the mechanism of activation of Sepharose by CNBr and alternative activation procedures, and lists numerous examples of proteins purified by affinity chromatography.

Contributed by Timothy A. Springer
Center for Blood Research
Harvard Medical School
Boston, Massachusetts

BCU IASI CENTRAL UNIVERSITY LIBRARY

Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography

Many biological processes, such as recombination, replication, and transcription, involve the action of sequence-specific DNA-binding proteins. Analysis and purification of these proteins by conventional chromatographic methods is often difficult because DNA-binding proteins typically make up <0.01% of the total cellular protein. Various methods have been developed for purifying such proteins. Affinity chromatography is a very effective means of purifying a protein based on its sequence-specific DNA-binding properties and is relatively straightforward if proper care is taken. The affinity chromatography procedure described in this unit uses DNA containing specific recognition sites for the desired protein that has been covalently linked to a solid support. Researchers in the past have purified proteins based on the proteins' ability to bind DNA, only to be disappointed to find that the DNA-binding activity of their highly purified products was not sequence-specific. This situation can be avoided by carefully performing all of the preliminary steps and experimental controls.

Basic Protocol 1 describes preparation of a DNA affinity resin, including cyanogen bromide (CNBr) activation of the agarose support. The Alternate Protocol provides a method to couple DNA to commercially available CNBr-activated Sepharose. Support Protocol 1 describes how to purify crude synthetic oligonucleotides by gel electrophoresis prior to preparation of the affinity resin. Basic Protocol 2 outlines the affinity chromatography procedure. Support Protocol 2 describes determination of the appropriate type and quantity of nonspecific competitor DNA that should be used in the procedure and its preparation. Parameters essential to the success of an affinity chromatography experiment are discussed in detail in the Commentary. Figure 9.6.1 provides a summary diagram of the entire chromatography procedure.

PREPARATION OF DNA AFFINITY RESIN

BASIC PROTOCOL 1

Correct choice of oligonucleotide sequence (discussed in detail in the Commentary) and preparation of the affinity resin are probably the most important parts of the affinity chromatography procedure. Preparation of affinity resin can be broken down into four steps: (1) preparing oligonucleotides; (2) activating Sepharose; (3) coupling DNA to resin; and (4) blocking unreacted CNBr. The first step requires highly purified oligonucleotides that contain the recognition sequence for the desired protein. Once purified, the complementary oligonucleotides are annealed, phosphorylated with T4 polynucleotide kinase, and ligated into long, multimeric chains (averaging ≥10-mers) with T4 DNA ligase. Sepharose CL-2B is activated with CNBr, the ligated DNA added, and the coupling reaction carried out overnight. Because CNBr is very toxic, researchers may prefer to employ the Alternate Protocol, which uses commercially available CNBr-activated Sepharose and therefore avoids direct handling of CNBr. After the coupling reaction, the remaining reactive groups are blocked with ethanolamine. This protocol is designed to prepare 10 ml of affinity resin.

NOTE: Glass-distilled or other high-quality water should be used throughout these procedures.

Contributed by Leslie A. Kerrigan and James T. Kadonaga
Current Protocols in Protein Science (1998) 9.6.1-9.6.18
Copyright © 1998 by John Wiley & Sons, Inc.

BCU LANCASTER UNIVERSITY LIBRARY

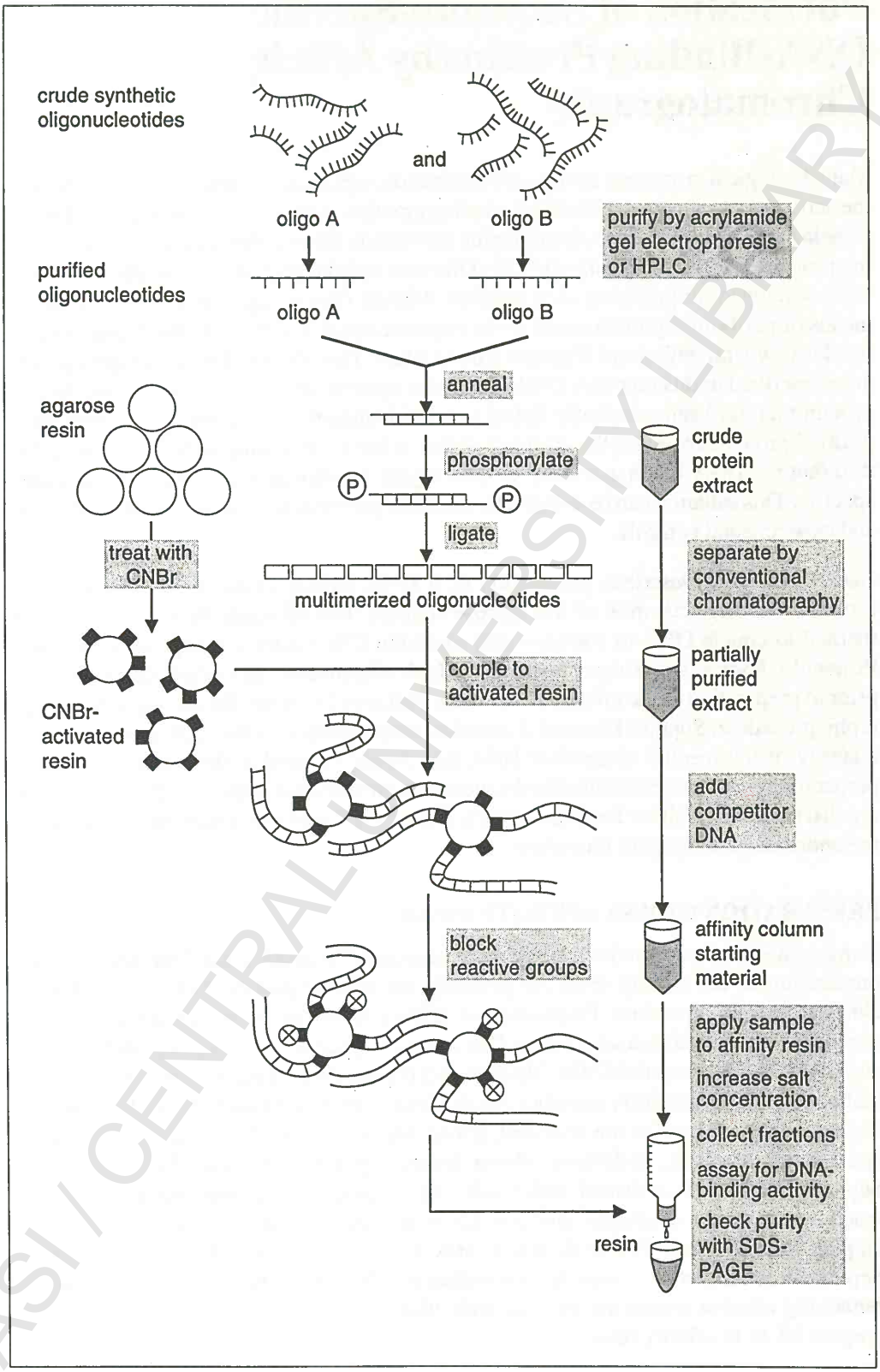


Figure 9.6.1 Purification of sequence-specific DNA-binding proteins with DNA affinity chromatography. Shown are the steps required to perform an affinity chromatography experiment using the methods described in this unit.

UCU IASI / CENTRAL UNIVERSITY LIBRARY

Materials

440 µg each of two synthetic oligonucleotides with desired binding site (see Support Protocol 1; or commercial HPLC-purified)
TE buffer, pH 7.8 (*APPENDIX 2E*)
10× T4 polynucleotide kinase buffer (see recipe)
20 mM ATP (Na^+ salt), pH 7.0
150 mCi/ml [γ-^{32}P]ATP (6000 Ci/mmol)
10 U/µl T4 polynucleotide kinase (New England Biolabs)
10 M ammonium acetate (*APPENDIX 2E*)
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol
24:1 (v/v) chloroform/isoamyl alcohol
3 M sodium acetate (*APPENDIX 2E*)
100% and 75% ethanol
10× linker/kinase buffer (see recipe)
6000 U/ml T4 DNA ligase (measured in Weiss units; New England Biolabs)
Buffered phenol
Isopropanol (2-propanol)
Sepharose CL-2B (Pharmacia Biotech)
Cyanogen bromide (CNBr; Aldrich)
N,*N*-dimethylformamide
5 N NaOH (*APPENDIX 2E*)
10 mM and 1 M potassium phosphate buffer, pH 8.0 (*APPENDIX 2E*)
1 M ethanolamine hydrochloride, pH 8.0 (see recipe)
NaOH, solid
Glycine
1 M KCl (*APPENDIX 2E*)
Column storage buffer (see recipe)

15-ml screw-cap polypropylene tubes
Heating blocks or water baths, 15°C, 37°C, 65°C, and 88°C
60-ml coarse-sintered glass funnel
Rotating wheel

5′-phosphorylate the oligonucleotides

1. In a 1.5-ml microcentrifuge tube, prepare a mixture containing 440 µg of each oligonucleotide in TE buffer in a total volume of 130 µl. Add 20 µl of 10× T4 polynucleotide kinase buffer. Incubate 2 min at 88°C, 10 min at 65°C, 10 min at 37°C, and 5 min at room temperature.

   A 1-µmol synthesis of oligonucleotide should yield enough purified DNA for ~20 ml of affinity resin.

2. Divide mixture in half in separate microcentrifuge tubes. To each 75-µl aliquot, add 15 µl of 20 mM ATP (pH 7.0), ~5 µCi [γ-^{32}P]ATP, and 10 µl of 10 U/µl T4 polynucleotide kinase (100 U total). Incubate 2 hr at 37°C.

   Dividing the reaction in half at this point facilitates the steps that follow. To add the labeled ATP, do not thaw, but simply touch (do not jab) the top of the frozen [γ-^{32}P]ATP with a yellow pipet tip and transfer the resulting tiny amount to the reaction tube.

3. Inactivate the kinase by adding 50 µl of 10 M ammonium acetate and 100 µl water to each tube and heating 15 min at 65°C. Allow to cool to room temperature.

BOSTON UNIVERSITY / CENTRAL LIBRARY

Purify phosphorylated oligonucleotides

4. Add 750 µl of 100% ethanol and mix by inversion. Microcentrifuge at high speed 15 min at room temperature to pellet DNA. Discard supernatant.
5. Resuspend each pellet in 225 µl TE buffer.
6. Add 250 µl of 25:24:1 phenol/chloroform/isoamyl alcohol to each tube. Vortex 1 min. Microcentrifuge 5 min at high speed to separate phases. Transfer the aqueous phase (upper layer) to a new tube.
7. Add 250 µl of 24:1 chloroform/isoamyl alcohol to aqueous phase. Vortex 1 min. Microcentrifuge 5 min at high speed to separate phases. Transfer the aqueous phase (upper layer) to a new tube.
8. Add 25 µl of 3 M sodium acetate to aqueous phase and mix by vortexing. Then add 750 µl of 100% ethanol and mix by inversion. Microcentrifuge 15 min at high speed to pellet DNA. Discard supernatant.
9. Wash pellet with 800 µl of 75% ethanol. Mix by vortexing. Microcentrifuge 5 min at high speed. Discard supernatant.
10. Dry pellet in vacuum evaporator (e.g., Speedvac).

Ligate oligonucleotides

11. Add 65 µl water and 10 µl of 10× linker/kinase buffer to each pellet. Dissolve DNA by vortexing. Add 20 µl of 20 mM ATP (pH 7.0) and 5 µl of 6000 U/ml T4 DNA ligase (30 Weiss units). Incubate ≥2 hr at room temperature or overnight at 15°C.

    Depending upon the oligonucleotides used, the optimal temperature for ligation will vary from 4° to 30°C. Short oligonucleotides (≤15-mers) tend to ligate better at lower temperatures (4° to 15°C), whereas oligonucleotides that have a moderate degree of palindromic symmetry tend to self-anneal and therefore ligate better at higher temperatures (15° to 30°C).

12. Monitor the ligation reaction by agarose gel electrophoresis, using 0.5 µl of ligation reaction per gel lane and including lanes containing size markers. Visualize DNA by ethidium bromide staining and UV photography.

    5′-phosphorylated oligonucleotides often do not ligate on the first attempt. If ligation has not occurred, extract the DNA once with 25:24:1 phenol/chloroform/isoamyl alcohol and once with 24:1 chloroform/isoamyl alcohol, then ethanol precipitate using sodium acetate as the precipitating salt. Dissolve the DNA in 225 µl TE, add 25 µl of 3 M sodium acetate, and reprecipitate with 750 µl of 100% ethanol. Wash with 75% ethanol, dry in Speedvac, and repeat ligation.

    The average length of the ligated oligonucleotides should be ≥10-mers. Ligating the oligonucleotides increases the binding capacity of the affinity resin and helps avoid potential problems due to steric hindrance between the factor and the Sepharose support. However, oligonucleotide multimers <10-mers will probably also work.

Purify oligonucleotide multimers

13. Add 100 µl buffered phenol to the 100-µl ligation reactions. Vortex 1 min. Microcentrifuge 5 min at high speed. Transfer aqueous phase (upper layer) to a new tube.
14. Add 100 µl of 24:1 chloroform/isoamyl alcohol to aqueous phase. Vortex 1 min. Microcentrifuge 5 min at high speed. Transfer aqueous phase (upper layer) to a new tube.
15. Add 33 µl of 10 M ammonium acetate to the aqueous phase. Mix by vortexing.

UCLA CENTRAL UNIVERSITY LIBRARY

16. Add 133 µl isopropanol. Mix by inversion. Incubate 20 min at –20°C. Microcentrifuge 15 min at high speed to pellet DNA. Discard supernatant.

The ammonium acetate/isopropanol precipitation removes residual ATP, which would otherwise interfere with coupling of the ligated DNA to the CNBr-activated Sepharose.

17. Add 225 µl TE buffer. Vortex to dissolve pellet. Add 25 µl of 3 M sodium acetate. Mix by vortexing. Add 750 µl of 100% ethanol. Mix by inversion. Microcentrifuge 15 min at high speed to pellet DNA. Discard supernatant.

18. Wash DNA twice with 75% ethanol. Dry pellet in vacuum evaporator.

19. Dissolve DNA in 50 µl water. Store at –20°C.

Do not dissolve the DNA in TE buffer, as the Tris buffer in TE will interfere with the coupling reaction.

Prepare CNBr-activated Sepharose

It is best to assemble all equipment and reagents required for the activation and coupling reactions *before* proceeding with the following steps.

20. Place 10 to 15 ml (settled bed volume) of Sepharose CL-2B in a 60-ml coarse-sintered glass funnel and wash extensively with 500 ml water.

To wash, add water to resin in funnel, stir gently with a glass rod, and remove water by vacuum suction, making sure to release vacuum before the resin is suctioned into a dry cake. Repeat until all 500 ml water is used.

21. Transfer moist Sepharose resin to a 25-ml graduated cylinder, estimating 10 ml of resin. Add water to 20 ml final volume. Transfer the resulting slurry to a 150-ml glass beaker containing a magnetic stir bar. Place beaker in a water bath equilibrated to 15°C and set up over a magnetic stirrer in a fume hood. Turn on stirrer to slow medium speed.

Keep the water bath at 15°C by periodically adding small chunks of ice as needed. Occasional variation of a degree or two is not critical.

22. In the fume hood, measure 1.1 g CNBr into a 25-ml Erlenmeyer flask, keeping the mouth of the flask covered with Parafilm or a ground glass stopper as much as possible (it is better to have slightly more than 1.1 g than slightly less). Add 2 ml *N,N*-dimethylformamide (the CNBr will dissolve instantly). Over the course of 1 min, add the resulting CNBr solution dropwise to the stirring Sepharose slurry.

CAUTION: *CNBr is highly toxic and volatile. Use only in a fume hood with extreme caution. Observe appropriate decontamination and disposal procedures.*

23. Immediately add 5 N NaOH as follows: add 30 µl to the stirring mixture every 10 sec for 10 min until 1.8 ml NaOH has been added.

It is convenient to measure 1.8 ml NaOH into a small tube before addition to the reaction to avoid having to change pipet tips during the 10-sec additions.

24. Immediately add 100 ml ice-cold water to the beaker and pour the mixture into a 60-ml coarse-sintered glass funnel.

IMPORTANT NOTE: *At this point, it is very important to avoid suction-filtering the resin into a dry cake. If the resin is accidentally dried into a cake, do not use; instead, repeat steps 20 to 24 with fresh resin.*

25. Still working in the fume hood, wash the resin in the funnel with four 100-ml washes of ice-cold (≤4°C) water followed by two 100-ml washes of ice-cold 10 mM potassium phosphate, pH 8.0.

BOSTON UNIVERSITY MEDICAL LIBRARY

26. Immediately transfer the resin to a 15-ml polypropylene screw-cap tube and add ~4 ml of 10 mM potassium phosphate (pH 8.0) until the resin has the consistency of a thick slurry.

Couple oligonucleotide multimers to CNBr-Sepharose

27. Immediately add the two 50-µl aliquots of DNA from step 19. Incubate on a rotating wheel overnight (≥8 hr) at room temperature.

28. In the fume hood, transfer the resin to a 60-ml coarse-sintered glass funnel and wash with two 100-ml washes of water and one 100-ml wash of 1 M ethanolamine hydrochloride, pH 8.0.

Using a Geiger counter, compare the level of radioactivity in the first few milliliters of filtrate with the level of radioactivity in the washed resin to estimate the efficiency of incorporation of DNA to the resin. Usually, all detectable radioactivity is present in the resin.

29. In the fume hood, transfer the resin to a 15-ml polypropylene screw-cap tube and add 1 M ethanolamine hydrochloride (pH 8.0) until the mixture is a smooth slurry. Incubate the tube on a rotating wheel 2 to 4 hr at room temperature.

This step inactivates unreacted CNBr-activated Sepharose.

It is important to clean up all CNBr waste carefully. In the fume hood, add solid NaOH and glycine (~10 to 20 mg/ml) to inactivate the CNBr. Soak contaminated instruments in a similar solution. Let sit overnight in the fume hood, then discard.

30. Wash the resin in a 60-ml coarse-sintered glass funnel with 100 ml of 10 mM potassium phosphate (pH 8.0), 100 ml of 1 M potassium phosphate (pH 8.0), 100 ml of 1 M KCl, 100 ml water, and 100 ml column storage buffer.

31. Store the resin at 4°C (stable at least 1 year; do not freeze).

ALTERNATE PROTOCOL

COUPLING THE DNA TO COMMERCIALLY AVAILABLE CNBr-ACTIVATED SEPHAROSE

The major advantage of this alternate procedure is that it begins with commercially available CNBr-activated chromatography resin, avoiding the need for preparation of CNBr-activated resin (see Basic Protocol 1). However, commercial CNBr-Sepharose is considerably more expensive than the homemade variety; moreover, the resulting column tends to run more slowly than one prepared as described above. Both resins are effective, leaving it up to the researcher's discretion which to use.

Additional Materials (also see Basic Protocol 1)

1 mM HCl (*APPENDIX 2E*), prepared fresh before use
CNBr-activated Sepharose 4B (Pharmacia Biotech)

1. Prepare oligonucleotide multimers by phosphorylation and ligation (see Basic Protocol 1, steps 1 to 19).

2. Weigh out 3 g CNBr-activated Sepharose 4B (1 g freeze-dried resin gives ~3.5 ml final gel volume).

3. Place the dry resin in a 15-ml conical polypropylene tube. Hydrate resin with 10 ml of 1 mM HCl and mix gently by flicking and inverting the tube. After 1 min, transfer slurry to a 60-ml coarse-sintered glass funnel. Wash and swell the beads by gradually pouring 500 ml of 1 mM HCl through the funnel (this will take ~15 min).

JCU IAS/CENTRAL UNIVERSITY LIBRARY

4. Wash the resin with 100 ml water and then with 100 ml of 10 mM potassium phosphate, pH 8.0.
5. Proceed with coupling of oligonucleotide multimers to the resin (see Basic Protocol 1, steps 26 to 31).

PURIFICATION OF OLIGONUCLEOTIDES BY PREPARATIVE GEL ELECTROPHORESIS

SUPPORT PROTOCOL 1

Preparation of affinity resin requires a large amount (1-µmol synthesis) of purified synthetic oligonucleotides. The oligonucleotides used to prepare the affinity resin must be of high purity because contaminating, incompletely synthesized oligonucleotides can interfere with the ligation reaction. HPLC-purified oligonucleotides are of sufficient purity for preparation of DNA affinity resins; however, it is normally expensive to obtain such oligonucleotides. An alternative is to purify crude oligonucleotides by the following method.

Additional Materials (also see Basic Protocol 1)

16% polyacrylamide-urea gel (see recipe)
Oligonucleotides to be purified
Formamide loading buffer (see recipe)
sec-butanol (2-butanol)
Diethyl ether
1 M $MgCl_2$ (*APPENDIX 2E*)

Saran wrap or other UV-transparent plastic wrap
Intensifying screen (e.g., Lightning Plus, NEN Life Sciences)
Hand-held short-wavelength UV light source
Silanized glass wool
Dry ice/ethanol bath (–78°C)

Additional reagents and equipment for denaturing polyacrylamide gel electrophoresis (*UNIT 10.3*)

Purify oligonucleotides by gel electrophoresis

1. Prepare a 20 cm × 40 cm × 1.5 mm denaturing polyacrylamide gel with four wells 3 cm in width, using 16% polyacrylamide/urea for separating oligonucleotides ~10 to ~45 bases long (or 8% or 6% polyacrylamide/urea for longer oligonucleotides). Let gel polymerize for ≥30 min, then prerun at 30 W for ≥1 hr.

2. Dissolve each oligonucleotide in formamide loading buffer to 200 µl final. Heat 15 min at 65°C to remove any secondary structure in the DNA. Load 50 µl of samples into separate wells, and run the gel at 30 W for ~4 hr.

   The amount of oligonucleotide that is prepared in a 1-µmol synthesis (~1 to 2 mg) can be loaded on one gel (0.25 µmol per 3-cm well). This is the maximum amount that can be applied to the gel without overloading it.

   It takes ~4 hr for the bromphenol blue in the loading buffet to migrate three-quarters of the way down the gel, which is far enough to purify oligonucleotides of 15 to 30 bases. In a 16% gel, bromphenol blue comigrates with ~10-base oligonucleotides and xylene cyanol comigrates with ~30-base oligonucleotides. If the oligonucleotide is 25 to 35 bases long, it is recommended that the formamide loading buffer be made without xylene cyanol.

Visualize gel by UV shadowing

3. Remove one of the glass gel plates and cover the gel with Saran wrap. Flip gel over and remove the other plate so that the gel is lying on the Saran wrap. Cover the other side of the gel with Saran wrap.

BOSTON COLLEGE CENTRAL UNIVERSITY LIBRARY

4. In a darkroom, lay the gel on an intensifying screen and hold a hand-held short-wavelength UV light source directly over it to visualize the DNA. Identify the major oligonucleotide band and mark its position directly on the Saran wrap using a marker, making sure that the light is directly over the band.

   The major band (usually the largest oligonucleotide but sometimes the next-to-largest) should be visible as a thick, dark band in the gel.

Purify oligonucleotides from gel

5. Carefully cut out the band with a razor blade, trying to avoid shredding the gel material or pulverizing the gel slice into small pieces.
6. Soak the gel piece in 5 ml TE buffer in a 15-ml polypropylene tube overnight at 37°C with shaking.
7. Place a silanized glass wool plug in a Pasteur pipet and prerinse with ~5 ml water. Filter the supernatant containing the DNA through the glass wool.
8. Concentrate the DNA to ≤180 µl by repeated extractions with *sec*-butanol.
9. Extract the DNA once with diethyl ether and place in a vacuum evaporator (e.g., Speedvac) until all traces of ether are removed.
10. Adjust the volume of the liquid to 180 µl with TE buffer. Add 20 µl of 3 M sodium acetate and 2 µl of 1 M $MgCl_2$. Mix by vortexing.
11. Add 600 µl of 100% ethanol. Mix by inversion. Chill 10 min in dry ice/ethanol. Let stand 5 min at room temperature. Microcentrifuge 15 min at high speed, room temperature, to pellet DNA. Discard supernatant.
12. Add 180 µl TE buffer. Dissolve pellet by vortexing. Add 20 µl of 3 M sodium acetate and 2 µl of 1 M $MgCl_2$. Mix by vortexing.
13. Add 600 µl of 100% ethanol. Mix by inversion. Chill 10 min in dry ice/ethanol. Let stand 5 min at room temperature. Microcentrifuge 15 min at high speed, room temperature. Discard supernatant.
14. Add 800 µl of 75% ethanol. Mix by vortexing. Microcentrifuge 5 min at high speed, room temperature. Discard supernatant.
15. Carefully dry pellet in vacuum evaporator (sometimes static electricity will cause the pellet to jump out of the tube).
16. Dissolve DNA in 200 µl TE buffer and measure A_{260} and A_{280}. Store at –20°C indefinitely.

    For sequence-specific DNA affinity resins, assume that 1 A_{260} unit = 40 µg/ml DNA.

BASIC PROTOCOL 2

DNA AFFINITY CHROMATOGRAPHY

Affinity chromatography is performed by combining a partially purified protein sample with appropriate competitor DNA, pelleting the insoluble protein-DNA complexes by centrifugation, and loading the resulting soluble material by gravity flow onto the affinity resin. Nonspecific DNA-binding proteins flow through the column while the specific protein is retained by the column. The protein is then eluted by gradually increasing the salt concentration of the buffer and individual fractions are tested for DNA-binding activity and purity. Fractions containing appropriate DNA-binding activity can be reapplied to the affinity resin if further purification is desired. By using two sequential affinity chromatography steps, a typical protein can be purified 500- to 1000-fold with ~30% yield. The method described below is performed with buffer Z, but many other buffers will work as well. Buffer choice is addressed in the Commentary.

IGU IAS/CENTRAL UNIVERSITY LIBRARY

Materials

Prepared DNA affinity resin (see Basic Protocol 1 or Alternate Protocol)
Buffer Z or other column buffer (e.g., buffers Z^e or TM; see recipes) made with varying KCl concentrations (buffers Z/0.1 M KCl through Z/1 M KCl)
Partially purified protein fraction dialyzed against buffer Z/0.1 M KCl
Nonspecific competitor DNA (see Support Protocol 2)
Column regeneration buffer (see recipe)
Column storage buffer (see recipe)

Disposable chromatography column (Poly-Prep, Bio-Rad)
Sorvall SS-34 rotor or equivalent
Liquid nitrogen
Narrow glass rod, silanized

Additional reagents and equipment for DNA-binding assays (Brenowitz, et al., 1989; Buratowski and Chodosh, 1996; Baldwin et al., 1996), SDS-PAGE (*UNIT 10.1*), and silver staining (*UNIT 10.5*)

1. Equilibrate 1 ml settled bed volume of the DNA affinity resin in a disposable chromatography column with two 10-ml washes of buffer Z/0.1 M KCl.
2. Combine the partially purified protein fraction in buffer Z/0.1 M KCl with nonspecific competitor DNA as determined by DNA-binding studies (see Commentary and Support Protocol 2).
3. Incubate mixture 10 min on ice.
4. Centrifuge mixture 10 min at 12,000 × *g* (10,000 rpm in Sorvall SS-34 rotor), 4°C, to pellet insoluble protein-DNA complexes.
5. Load the supernatant onto the column at gravity flow (e.g., 15 ml/hr per column for Sepharose CL-2B).

   A single 1-ml column is sufficient for a standard nuclear extract (e.g., 12 liters of HeLa cells or 150 g of Drosophila embryos). When purifying a larger quantity of material, it is preferable to use multiple 1-ml columns. It is common practice to apply as much as 50 ml of protein sample onto a single 1-ml column.

   Typically, 1 ml of affinity resin contains ~80 to 90 µg DNA, which corresponds to a protein-binding capacity of 7 nmol/ml of resin, assuming one recognition site per 20 bp.

6. After loading the starting material, wash the column four times with 2-ml aliquots of buffer Z/0.1 M KCl, rinsing the sides of the column each time.

   It is very important to wash the affinity column thoroughly at this step by rinsing down the sides of the column with the 2-ml wash buffer aliquots; a single 8-ml wash is not effective.

   DNA affinity columns will often yield >100-fold purification of the desired factor. When a protein is purified 100-fold, however, a 1% contamination of the starting material due to inefficient washing of the column will lead to major contamination of the affinity-purified factor.

7. Elute the protein from the column by successive addition of 1-ml portions of buffer Z/0.2 M KCl, buffer Z/0.3 M KCl, buffer Z/0.4 M KCl, buffer Z/0.5 M KCl, buffer Z/0.6 M KCl, buffer Z/0.7 M KCl, buffer Z/0.8 M KCl, and buffer Z/0.9 M KCl, followed by three 1-ml aliquots of buffer Z/1 M KCl. Collect 1-ml fractions that correspond to the addition of the 1-ml portions of buffer. Quick-freeze the protein samples in liquid nitrogen and store at –80°C. Samples can generally be stored for at least 2 years.

B.C.U. „M. Eminescu" IAŞI
CHIMIE

BOSTON UNIVERSITY LIBRARY

It is convenient to save separate aliquots of each fraction for DNA-binding assays (20 µl each) and SDS-PAGE analysis (50 µl each).

8. Assay the protein fractions for the sequence-specific DNA-binding activity using a DNA-binding assay (Brenowitz, et al., 1989; Buratowski and Chodosh, 1996; Baldwin et al., 1996). Estimate the purity of the protein fractions by SDS-PAGE (*UNIT 10.1*) followed by silver staining to visualize the protein.

   If further purification is desired, combine the fractions that contain the activity and, depending on the KCl concentration, either dilute (using buffer Z without KCl) or dialyze (against buffer Z/0.1 M KCl) to 0.1 M KCl. Then combine the protein fraction with nonspecific competitor DNA (the amount and type of which must be determined experimentally as described in Support Protocol 2) and reapply to either fresh or regenerated DNA affinity resin.

   In some instances, proteins might still be bound to the affinity resin after the 1 M salt elution step (step 7). Thus, if the desired protein is not detected in either the column fractions or flowthrough, it is possible that the protein is still on the column, and a wash at a higher salt concentration (e.g., 2 M KCl) may be needed to elute the protein from the resin.

9. Regenerate the affinity resin as follows: At room temperature, stop the column flow and add 5 ml column regeneration buffer to the column. Stir the resin with a silanized narrow glass rod to mix the resin with the regeneration buffer. Let the buffer flow out of the column. Repeat step.

10. To store column, add 10 ml column storage buffer and allow to flow through. Repeat this wash, then close the bottom of the column and add another 5 ml buffer. Cover top of column and store at 4°C.

    The column may be stored for ≤1 year. Alternatively, the affinity resin may be removed from the column after the two washes and stored in a clean polypropylene tube with the 5 ml buffer.

SUPPORT PROTOCOL 2

SELECTION AND PREPARATION OF NONSPECIFIC COMPETITOR DNA

Proper use of competitor DNA is essential for successful purification of DNA-binding proteins by affinity chromatography. Common competitor DNAs include poly(dI-dC), poly(dA-dT), poly(dG-dC), and calf thymus DNA. The amount and type of competitor to use in a given experiment must be determined experimentally. A typical method is to mix a protein sample (the exact fraction that will be applied to the affinity resin) with varying amounts of different competitor DNAs and evaluate DNA binding using assays such as DNase I footprinting (Brenowitz et al., 1989) or gel mobility shifts (Buratowski and Chodosh, 1996). The competitor DNA that inhibits DNA binding the least should be used for DNA affinity chromatography. It is perhaps easiest to think of these binding experiments as scaled-down affinity columns. First, determine the highest amount of competitor DNA that can be added to a DNA-binding reaction that does not interfere with the binding of the sequence-specific factor. To move up to a full-scale experiment, use one-fifth of the amount that would be required if the binding reaction were directly scaled up.

The following example with a hypothetical factor M illustrates how to determine the amount of competitor to use in an affinity chromatography experiment. Using DNase I footprinting, a strong footprint was observed with 5 µl of a 0.4 M heparin fraction. Testing various competitors reveals that calf thymus DNA, poly(dI-dC), and poly(dG-dC) all strongly inhibit binding of factor M. However, no detectable inhibition of factor M binding is observed with 2 µg poly(dA-dT), weak but detectable inhibition is observed with 3 µg poly(dA-dT), and strong inhibition is observed with 4 µg poly(dA-dT). In this example,

UCLA CENTRAL UNIVERSITY LIBRARY

with 5 µl of the 0.4 M heparin fraction, the highest amount of poly(dA-dT) that does not inhibit factor M binding is 2 µg. If 5 ml of the 0.4 M heparin fraction is available, the direct scale-up would be 1000-fold, so 1000 × 2 µg = 2000 µg poly(dA-dT) would be needed. Because in an affinity chromatography experiment one-fifth of the direct scale-up amount is used, the appropriate amount of poly(dA-dT) to add to 5 ml of the 0.4 M heparin fraction is 2000 µg × 1/5 = 400 µg.

The optimal amount of competitor DNA will vary with the purity of the partially purified protein sample. Fractions that contain more nonspecific DNA-binding proteins will normally require more competitor DNA and thus it is necessary to determine experimentally the optimal amount of DNA to use with each protein fraction. In addition, the amount of competitor DNA is estimated as the mass of DNA (in micrograms) to add per volume (in milliliters) of protein fraction; this is because it is likely that competitor DNA acts by forming a complex with high-affinity, nonspecific DNA-binding proteins in the crude extract. Different competitor DNAs can be used in a single experiment.

To prepare poly(dI-dC), poly(dG-dC), and poly(dA-dT), dissolve the desired amount of competitor DNA to a final concentration of 10 A_{260} units in TE buffer per 100 mM NaCl. Heat the sample to 90°C and slowly cool to room temperature over 30 to 60 min. If the average length of the DNA is >1 kb, degrade it by sonication. Estimate the length of the DNA by agarose gel electrophoresis.

REAGENTS AND SOLUTIONS

Buffer TM

50 mM Tris·Cl, pH 7.9 (*APPENDIX 2E*)
0 M *or* 1 M KCl
12.5 mM $MgCl_2$
1 mM dithiothreitol (DTT; add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) Nonidet P-40 (NP-40)

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively, and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

Buffer Z

25 mM HEPES (K^+ salt), pH 7.6
0 M *or* 1 M KCl
12.5 mM $MgCl_2$
1 mM DTT (add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) NP-40
Adjust pH to 7.6 with KOH

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively, and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

BCU RAY CENTRAL UNIVERSITY LIBRARY

Buffer Z[e]

25 mM HEPES (K^+ salt), pH 7.6
0 M *or* 1 M KCl
1 mM DTT (add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) NP-40
Adjust the pH to 7.6 with KOH

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively, and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

Column regeneration buffer

10 mM Tris·Cl, pH 7.8 (*APPENDIX 2E*)
1 mM EDTA, pH 8.0
2.5 M NaCl
1% (v/v) NP-40
Store at room temperature

The solution will be cloudy and separate into two phases (NP-40 and aqueous) upon storage. Mix by swirling and shaking just before use.

Column storage buffer

10 mM Tris·Cl, pH 7.8 (*APPENDIX 2E*)
1 mM EDTA, pH 8.0
0.3 M NaCl
0.04% (w/v) sodium azide

Store at room temperature without sodium azide. Make a 4% (w/v) sodium azide stock solution and add just before use.

Ethanolamine hydrochloride, pH 8.0, 1 M

1 M ethanolamine
Adjust pH to 8.0 with HCl
Filter sterilize
Store at room temperature

Formamide loading buffer

90 ml deionized formamide
10 ml 10× TBE (see recipe below)
40 mg xylene cyanol
40 mg bromphenol blue
Store at –20°C

Linker-kinase buffer, 10×

660 mM Tris·Cl, pH 7.6 (*APPENDIX 2E*)
100 mM $MgCl_2$
100 mM DTT
10 mM spermidine
Store at –20°C
Add an extra 10 mM DTT just before use

DCU IAST / CENTRAL UNIVERSITY LIBRARY

16% polyacrylamide-urea gel

50 ml 40% (w/v) 19:1 acrylamide/bisacrylamide
12.5 ml 10× TBE (see recipe below)
62.5 g urea
17 ml H_2O
Mix, filter, briefly degas, and then add:
750 µl 10% (w/v) ammonium persulfate
20 µl TEMED

Pour immediately into prepared gel plates.

T4 polynucleotide kinase buffer, 10×

500 mM Tris·Cl, pH 7.6 (*APPENDIX 2E*)
100 mM $MgCl_2$
50 mM DTT
1 mM spermidine
1 mM EDTA, pH 8.0
Store at –20°C
Add an extra 50 mM DTT just before use

TBE (Tris/borate/EDTA) electrophoresis buffer, 10×

108 g Tris base (890 mM)
55 g boric acid (890 mM)
900 ml H_2O
40 ml 0.5 M EDTA, pH 8.0 (20 mM)
H_20 to 1 liter

COMMENTARY

Background Information

Purification of sequence-specific DNA-binding proteins has historically been a difficult task, mainly because such proteins are a small fraction of the total cellular protein. However, it is now possible to purify these factors quickly, simply, and effectively by using multimerized synthetic oligonucleotides that contain the recognition sequence for a particular DNA-binding protein. Early efforts with DNA-affinity chromatography involved adsorption or coupling of nonspecific DNA (such as calf thymus DNA) to either cellulose (Alberts and Herrick, 1971) or agarose (Arndt-Jovin et al., 1975) supports. These methods paved the way for the development of a variety of sequence-specific DNA affinity chromatography techniques, including the procedures described in this unit.

Other methods have been described to purify sequence-specific DNA-binding proteins, including chromatography using biotinylated DNA fragments attached to various supports by biotin-avidin or biotin-streptavidin coupling (*UNIT 10.6*; Chodosh et al., 1986; Kasher et al., 1986; Leblond-Francillard et al., 1987), oligonucleotides synthesized onto Teflon-based beads (Duncan and Cavalier, 1988), or synthetic oligonucleotide monomers attached to agarose supports (Wu et al., 1987; Blanks and McLaughlin, 1988; Hoey et al., 1993); and preparative gel mobility shifts (Gander et al., 1988). Although more than 50 sequence-specific DNA-binding proteins have been purified by the method described in this unit (Kadonaga, 1991, and references therein), it is likely that many of the techniques listed above are also effective for purifying sequence-specific DNA-binding proteins.

There are a variety of reasons that CNBr activation is commonly used in the preparation of affinity resins: it is simple, works well with agarose matrices, and is mild enough to bind ligands such as DNA. Briefly, the chemistry of the CNBr activation reaction is as follows. At high pH, the hydroxyl groups on the agarose resin react with CNBr. The majority of the CNBr added to the reaction reacts with water to yield inert cyanate ions, which is part of the reason such a large amount of CNBr is required. Additionally, the majority of the cyanate esters that are formed on the agarose either are hydrolyzed to form inert carbamate or react with the matrix hydroxyls to form imidocarbonates. The imidocarbonates that form can act effectively

BOSTON UNIVERSITY LIBRARY

as chemical cross-links, thus stabilizing the matrix (which is in most cases beneficial, particularly if the agarose resin chosen is not covalently cross-linked). The remaining active cyanate esters are coupled to the amino-containing ligands (in this case, oligonucleotides) at physiological pH. Finally, the unreacted cyanate esters are blocked with an excess of a suitable reagent, such as ethanolamine, to prevent coupling of the protein sample to the matrix (Janson and Rydén, 1989).

Critical Parameters and Troubleshooting

Basic strategy

The general approach to affinity purification of DNA-binding proteins is as follows. First, estimate the binding site of the desired protein by a technique such as DNase I footprinting (Brenowitz et al., 1989). If possible, it is best to survey a variety of promoters and enhancers. Next, determine the optimal conditions for protein binding to the DNA, considering factors such as temperature, ionic strength, pH, and Mg^{2+} concentration. Using conventional chromatography, partially purify the protein to remove contaminants such as nucleases that may degrade the affinity resin. Test a variety of nonspecific competitor DNAs, including poly(dI-dC), poly(dG-dC), poly(dA-dT), and calf thymus DNA to determine their effect on protein-DNA interactions. For best results, prepare two or more different DNA affinity resins with naturally occurring, high-affinity binding sites, preferably containing different flanking DNA sequences. The desired protein should bind with high affinity to both resins; if it does not, then it is likely that the protein that has been purified is not specific to the desired binding site. Finally, it is important to prepare a control resin that does not contain the recognition sequence for the desired protein. A control resin will make it possible to identify proteins that bind nonspecifically to DNA-Sepharose. Using this approach, it is possible to obtain a preparation containing a highly purified, sequence-specific DNA-binding protein.

Starting material and conventional chromatography

Extensive purification of the DNA-binding protein is not required for effective affinity chromatography. However, partial purification of the protein by conventional chromatography is recommended prior to affinity chromatography to remove proteases and nucleases that might degrade either the protein or the affinity resin. Keep in mind that sequence-specific and nonspecific DNA-binding proteins will often copurify in conventional chromatography, possibly because they may both interact with negatively charged polymers that resemble DNA. It is important to note that the most persistent contaminants in sequence-specific DNA-binding protein preparations are nonspecific, high-affinity DNA-binding proteins. Various conventional chromatographic methods may be employed prior to affinity purification of a DNA-binding protein; methods that have been used successfully in the past are discussed below.

Ion-exchange chromatography. One obvious first step might be cation-exchange chromatography with resins such as S-Sepharose Fast Flow/Mono S (Pharmacia Biotech), CM-52 (Whatman), CM Sepharose Fast Flow (Pharmacia Biotech), P11 phosphocellulose (Whatman), or Bio-Rex 70 (Bio-Rad). DNA-binding proteins (both specific and nonspecific) will normally bind to cation-exchange resins in buffers containing from 50 mM to 100 mM NaCl, and can be eluted with higher salt concentrations.

Affinity chromatography. Nonspecific DNA-cellulose (Alberts and Herrick, 1971) or DNA-agarose (Arndt-Jovin et al., 1975) resins can be used for a preliminary purification step. These resins are typically prepared with salmon sperm or calf thymus DNA. Also, sequence-specific DNA affinity resins can be prepared with oligonucleotides that do not contain binding sites for the desired factor. A strategy for chromatography with either nonspecific DNA affinity resin (e.g., see Rosenfeld and Kelly, 1986) or sequence-specific DNA affinity resins that lack the binding site for the desired factor (e.g., see Kaufman et al., 1989) is as follows. Once the protein sample is applied to a nonspecific DNA affinity resin, the desired factor will either (1) flow through the resin or (2) elute from the resin at a low salt concentration. High-affinity, nonspecific DNA-binding proteins should remain bound to the resin and will thus be separated from the desired protein. If this strategy is successful, it will not be necessary to add the competitor DNA prior to the subsequent sequence-specific affinity chromatography step because the high-affinity nonspecific DNA-binding proteins will already have been separated from the sample.

A resin that works well if the desired protein contains O-linked *N*-acetylglucosamine monosaccharide residues is wheat germ agglutinin-agarose. This method has been used successfully in the past with transcription factors such as Sp1 (Jackson and Tjian, 1989) and HNF1 (Lichtsteiner and Schibler, 1989).

Heparin-agarose chromatography. Heparin-agarose resins are also excellent for the purification of DNA-binding proteins because they possess properties similar to those of both ion-exchange and affinity resins. It is important to be aware that variability exists between different heparin-agarose preparations. This variability may be due to differences in the methods of coupling heparin to agarose, because alternate methods for coupling and blocking the agarose may produce different functional groups on the resin. If a particular batch of heparin-agarose is used successfully, it is wise to continue to use that batch to ensure reproducibility.

Gel-filtration chromatography. Finally, gel-filtration chromatography may be used effectively for partial purification of a DNA-binding protein because separation is based on size and shape, not DNA-binding properties. For example, gel filtration may be useful for separating a desired protein from contaminating nucleases and other nonspecific DNA-binding proteins, in contrast to chromatography using ion-exchange, heparin-agarose, and nonspecific DNA affinity resins that may actually enrich for DNA-binding proteins. In some cases gel filtration may be undesirable because significant sample dilution occurs; however, further concentration can be obtained by subsequent affinity chromatography.

DNA binding studies

Binding conditions. In order to use affinity chromatography to purify a DNA-binding protein, it is necessary to optimize binding conditions. It is important to determine these conditions experimentally, as ionic strength, temperature, pH, and presence or absence of Mg^{2+} can affect protein-DNA interactions. In addition, although HEPES and PIPES are often considered to be superior to Tris buffers, some commercially available preparations of HEPES and PIPES buffers contain contaminants that inhibit binding of proteins to DNA. Therefore, in some respects, it may be safer to use Tris buffer. It is also important to note that some factors, such as Sp1, bind with higher affinity to DNA in the absence of Mg^{2+} than in its presence. Thus, in some cases it may be useful to omit Mg^{2+} from the chromatography buffer. In addition, although the procedures in this unit use buffers containing KCl, NaCl could probably be substituted. Although it is usually preferable to handle proteins at 4°C, proteins have been described that will not bind DNA at 4°C, yet bind perfectly well at room temperature; in such cases the temperature at which the affinity chromatography is performed should be adjusted accordingly. Finally, to minimize nonspecific adsorption of affinity-purified proteins to plastic and glass, chromatography and storage buffers should contain a nonionic detergent such as NP-40 at a concentration of ~0.01% to 0.1% (v/v).

Binding assays. There are many methods for identifying proteins that are bound to DNA, including DNAse I footprinting (Galas and Schmitz, 1978; Brenowitz et al., 1989), Fe-EDTA footprinting (Tullius et al., 1987), methidiumpropyl-EDTA·Fe^{2+} (MPE; Hertzberg and Dervan, 1982), methylation interference (Baldwin et al., 1996), and gel mobility shift assays (Buratowski and Chodosh, 1996). Although binding information may be obtained from all these techniques, it is most straightforward to use DNase I footprinting to identify a protein binding site. Additionally, DNase I is active over a broad range of buffer conditions, which allows for considerable variation in binding conditions. Visualization of the protected region by footprinting makes designing oligonucleotides for affinity chromatography relatively simple (as discussed later in this section). Both DNase I footprinting and gel mobility shift assays are useful for assaying affinity-column fractions.

Design of oligonucleotides

Many different parameters are important in designing oligonucleotides for an affinity resin. The first step, as described above, is to determine the protein-binding site on the DNA. The DNA affinity resin should be prepared from naturally occurring, high-affinity binding sites. Using a consensus sequence may work in some instances, but in most situations a "real" binding site is best. Given that the length of the oligonucleotides can be reasonably flexible (14-mers to 61-mers have been used successfully), use of oligonucleotides that include the DNase I footprint as well as some flanking DNA (at least a few bases beyond the borders of the footprint) is recommended. In the early stages, it is better to copurify factors that bind

BOSTON CENTRAL UNIVERSITY LIBRARY

to the flanking region than to fail to purify the desired factor because the binding site is too short. It may be wise to avoid using oligonucleotides that are 21 or 42 bases in length because the DNA, if bent, might have a greater tendency to circularize during the ligation step. The oligonucleotides should also be designed with a single-stranded overhang, for two reasons: first, the ligation reaction will be more efficient (as compared to a blunt-end ligation), and second, it is likely that the DNA couples to the resin via the primary amine groups of the single-stranded overhang of the ligated multimers. The sequence GATC-XXX-BINDING SITE-XXX (where XXX represents the DNA flanking the binding site) works well, though it may also be convenient to use the sequences flanking the binding site as the overhang.

Inadvertent purification of nonspecific DNA-binding proteins

Throughout this unit, an attempt has been made to emphasize how to avoid inadvertent purification of nonspecific DNA-binding proteins. Unfortunately, this problem is quite common, particularly when the gel mobility shift assay is used to monitor DNA binding. Proper use of the control resin (to identify proteins that bind nonspecifically to DNA-Sepharose), two different affinity resins (to identify proteins that might be interacting with the linker DNA), and competitor DNA (to deplete nonspecific DNA-binding proteins from the extract) should virtually eliminate purification of nonspecific DNA-binding proteins. In addition, DNase I footprinting is recommended for assaying DNA binding because it allows visualization of the actual protein binding site and to confirm sequence specificity. To correlate DNA binding with a particular polypeptide, it would be informative to purify the polypeptide from a polyacrylamide gel by using the denaturation/renaturation procedure described by Hager and Burgess (1980). Two common nonspecific DNA-binding proteins that have been purified from HeLa cells by affinity chromatography are poly(ADP-ribose) polymerase, which has an M_r of 116,100 (Ueda and Hayaishi, 1985; Slattery et al., 1983), and the Ku antigen, which consists of two polypeptides of M_r 70,000 and 80,000 (Mimori et al., 1986).

Anticipated Results

After two rounds of affinity chromatography, it is possible to achieve a 500- to 1000-fold purification with a 30% overall yield. The concentration of affinity-purified proteins is typically 5 to 50 µg/ml. About 0.2 to 1 µl of protein (2 to 10 ng) can be expected to be sufficient for a "blanked-out" footprint (with ~10 fmol of DNA probe). Affinity resin is stable for >1 year if stored at 4°C. Affinity-purified proteins are typically stable for >2 years if stored at –80°C in an appropriate buffer and handled properly. Proteins should always be quick-frozen in liquid nitrogen and quick-thawed in either cold or room-temperature water (although if room-temperature water is used to accelerate thawing, the sample must not be allowed to warm up above 4°C). Although many DNA-binding proteins remain active after several freeze-thaw cycles, it is preferable to divide protein preparations into small aliquots rather than to freeze and thaw the entire sample many times.

Time Considerations

The time required to optimize DNA-binding conditions of a particular protein is variable and depends on assay type, probe preparation time, and competitor choices. The time to prepare extracts and to perform conventional chromatography is also variable depending on factors such as starting material and type of chromatography. Beginning with purified oligonucleotides, preparation of the DNA affinity resin takes 2 days (with either commercially available or homemade CNBr-activated resin). If the oligonucleotides are to be gel-purified, an additional day and a half should be added. Affinity chromatography takes ~4 to 8 hr, dependent mainly on the flow rate of the resin: Sepharose CL-2B runs at ~15 ml/hr, while Sepharose-4B runs somewhat slower. Therefore, it is possible to estimate chromatography time by calculating the volume of sample that is to be applied to the column plus 8 ml for the wash steps and 11 ml for the elution steps. If the volume of starting material is not too large, it is possible to assay the affinity column on the day it is run (either by DNA-binding assays, SDS-PAGE, or both). It is important to freeze the affinity fractions in liquid nitrogen and store at –80°C as soon as the column is complete.

Literature Cited

Alberts, B. and Herrick, G. 1971. DNA-cellulose chromatography. *Methods Enzymol.* 21:198-217.

Arndt-Jovin, D.J., Jovin, T.M., Bähr, W., Frischauf, A.-M., and Marquardt, M. 1975. Covalent attachment of DNA to agarose: Improved synthesis and use in affinity chromatography. *Eur. J. Biochem.* 54:411-418.

BCU IASCHER INSTITUTE LIBRARY

Baldwin, A.S., Oettinger, M., and Struhl, K. 1996. Methylation and uracil interference assays for analysis of protein-DNA interactions. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 12.3.1-12.3.7. John Wiley & Sons, New York.

Blanks, R. and McLaughlin, L.W. 1988. An oligodeoxynucleotide affinity column for the isolation of sequence specific DNA binding proteins. *Nucl. Acids Res.* 16:10283-10299.

Brenowitz, M., Senear, D.F., and Kingston, R.E. 1989. DNase I footprint analysis of protein-DNA binding. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 12.4.1-12.4.16. John Wiley & Sons, New York.

Buratowski, S. and Chodosh, L. 1996. Mobility-shift DNA-binding assay using gel electrophoresis. *In* Current Protocols in Molecular Biology. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 12.2.1-12.2.11. John Wiley & Sons, New York.

Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and transcription activities of the *Adenovirus* major late transcription factor. *Mol. Cell. Biol.* 6:4723-4733.

Duncan, C.H. and Cavalier, S.L. 1988. Affinity chromatography of a sequence-specific DNA binding protein using Teflon-linked oligonucleotides. *Anal. Biochem.* 169:104-108.

Galas, D. and Schmitz, A. 1978. DNase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucl. Acids Res.* 5:3157-3170.

Gander, I., Foeckler, R., Rogge, L., Meisterernst, M., Schneider, R., Mertz, R., Lottspeich, F., and Winnacker, E.L. 1988. Purification methods for the sequence-specific DNA-binding protein nuclear factor I (NFI)—generation of protein sequence information. *Biochim. Biophys. Acta* 951:411-418.

Hager, D.A. and Burgess, R.R. 1980. Elution of proteins from sodium dodecyl sulfate polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: Results with sigma subunit of *Escherichia coli* RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. *Anal. Biochem.* 109:76-86.

Hertzberg, R.P. and Dervan, P.B. 1982. Cleavage of double helical DNA by (methidiumpropyl-EDTA)iron(II). *J. Am. Chem. Soc.* 104:313-315.

Hoey, T., Welnzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D., and Tjian, R. 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* 72:247-260.

Jackson, S.P. and Tjian, R. 1989. Purification and analysis of RNA polymerase II transcription factors using wheat germ agglutinin affinity chromatography. *Proc. Natl. Acad. Sci. U.S.A.* 86:1781-1785.

Janson, J.-C. and Rydén, L. 1989. Protein Purification. Principles, High Resolution Methods, and Applications. VCH Publishers, New York.

Kadonaga, J.T. 1991. Purification of sequence-specific DNA binding proteins by DNA affinity chromatography. *Methods Enzymol.* 208:10-23.

Kasher, M.S., Pintel, D., and Ward, D.C. 1986. Rapid enrichment of HeLa transcription factors IIIB and IIIC by using affinity chromatography based on avidin-biotin interactions. *Mol. Cell. Biol.* 6:3117-3127.

Kaufman, P.D., Doll, R.F., and Rio, D.S. 1989. *Drosophila* P element transposase recognizes internal P element DNA sequences. *Cell* 59:359-371.

Leblond-Francillard, M., Dreyfus, M., and Rougeon, F. 1987. Isolation of DNA-protein complexes based on streptavidin and biotin interaction. *Eur. J. Biochem.* 166:351-355.

Lichtsteiner, S. and Schibler, U. 1989. A glycosylated liver-specific transcription factor stimulates transcription of the albumin gene. *Cell* 57:1179-1187.

Mimori, T., Hardin, J.A., and Steitz, J.A. 1986. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J. Biol. Chem.* 261:2274-2278.

Rosenfeld, P.J. and Kelly, T.J. 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261:1398-1408.

Slattery, E., Dignam, J.D., Matsui, T., and Roeder, R.G. 1983. Purification and analysis of a factor which suppresses nick-induced transcription by RNA polymerase II and its identity with poly(ADP-ribose) polymerase. *J. Biol. Chem.* 258:5955-5959.

Tullius, T.D., Dombroski, B.A., Churchill, M.E.A., and Kam, L. 1987. Hydroxyl radical footprinting: A high resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155:537-558.

Ueda, K. and Hayaishi, O. 1985. ADP-ribosylation. *Annu. Rev. Biochem.* 54:73-100.

Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V., and Ueda, H. 1987. Purification and properties of *Drosophila* heat shock activator protein. *Science* 238:1247-1253.

Key References

Kadonaga, J.T. 1991. See above.

Techniques paper, though less descriptive than this unit, containing a table that lists (with references) >50 sequence-specific proteins that have been purified using the affinity chromatography method described herein.

BOWMAN GRAY SCHOOL OF MEDICINE LIBRARY

Kadonaga, J.T. and Tjian, R. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 83:5889-5893.

First paper to describe affinity chromatography with multimerized oligonucleotides; details purification of transcription factor Sp1.

Contributed by Leslie A. Kerrigan
Osiris Therapeutics
Baltimore, Maryland

James T. Kadonaga
University of California San Diego
La Jolla, California

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Purification of DNA-Binding Proteins Using Biotin/Streptavidin Affinity Systems

Short fragments of DNA—either natural or formed from oligonucleotides—containing a high-affinity site for a DNA-binding protein provide a powerful tool for purification. The biotin/streptavidin purification system is based on the tight and essentially irreversible complex that biotin forms with streptavidin. The experimental design of this system is illustrated in Figure 9.7.1. First, a DNA fragment is prepared that contains a high-affinity binding site for the protein of interest. A molecule of biotinylated nucleotide is incorporated into one of the ends of the DNA fragment. The protein of interest is allowed to bind to the high-affinity recognition site present in the biotinylated fragment. The tetrameric protein streptavidin is then bound to the biotinylated end of the DNA fragment. Next, the protein/biotinylated fragment/streptavidin ternary complex is efficiently removed by adsorption onto a biotin-containing resin. Since streptavidin is multivalent, it is able to serve as a bridge between the biotinylated DNA fragment and the biotin-containing resin. Proteins remaining in the supernatant are washed away under conditions that maximize the stability of the DNA-protein complex. Finally, the protein of interest is eluted from the resin with a high-salt buffer.

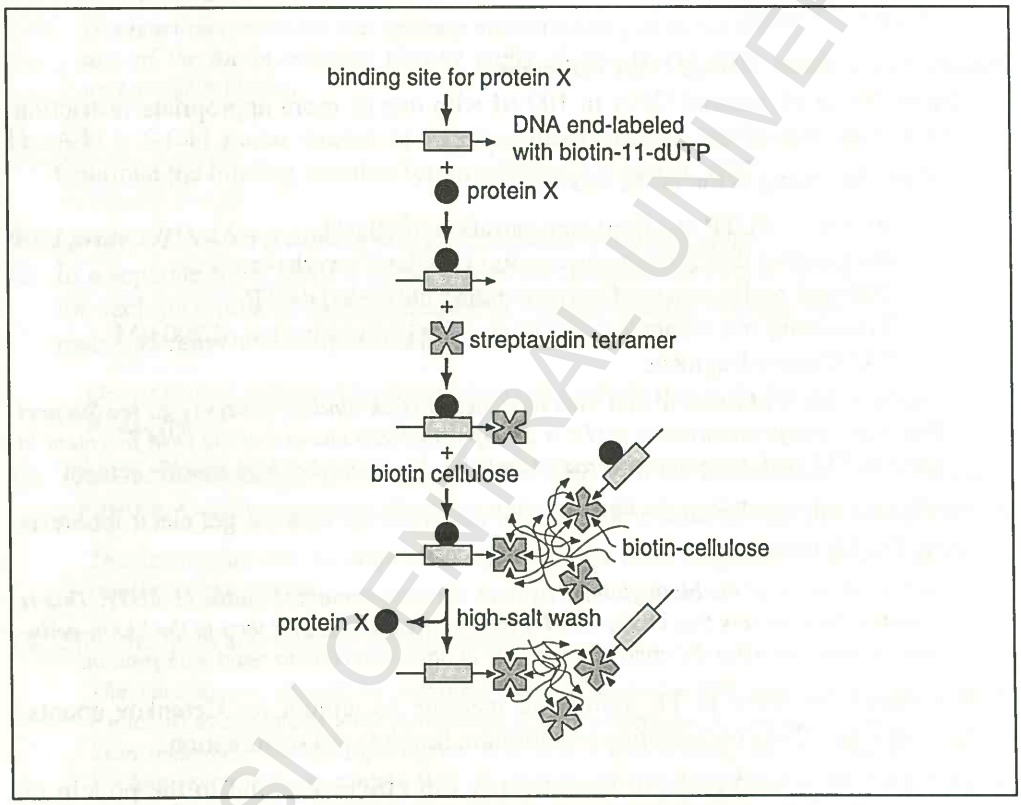


Figure 9.7.1 Purification of DNA-binding proteins using the biotin/streptavidin affinity technique. The protocol involves the following steps: (1) a biotinylated, labeled DNA fragment is prepared containing a binding site for the protein to be purified; (2) the biotin-cellulose resin is prepared; (3) a binding reaction containing a crude protein fraction and the biotinylated probe is set up; (4) free streptavidin is added to the binding reaction; (5) the protein/biotinylated DNA fragment/streptavidin complex is bound to the biotin-cellulose resin; (6) unbound protein is removed by extensive resin washing; and (7) the protein is eluted from the resin with high-ionic-strength buffer. Each of these steps can be monitored and optimized in solution, using a mobility shift DNA-binding assay (see Support Protocol and Buratowski and Chodosh, 1996).

Contributed by Lewis A. Chodosh and Stephen Buratowski
Current Protocols in Protein Science (1998) 9.7.1-9.7.13
Copyright © 1998 by John Wiley & Sons, Inc.

BIBLIOTECA CENTRAL UNIVERSITY LIBRARY

PURIFICATION USING BATCH METHOD

Materials

Plasmid DNA with binding site for the protein of interest
Appropriate restriction endonucleases
Biotin-11-dUTP (Life Technologies)
Labeled and unlabeled dNTPs
Klenow fragment of *E. coli* DNA polymerase I
TE buffer (*APPENDIX 2E*)
Biotin-cellulose (Pierce)
Biotin-cellulose binding buffer (see recipe)
BSA
Bulk carrier DNA [e.g., poly(dI-dC)·poly(dI-dC), salmon sperm DNA, or *E. coli* DNA]
Biotin-cellulose elution buffer (see recipe)
Protein solution
Streptavidin (Celltech; may be stored as 5 mg/ml stock for at least 2 months)

DEAE membrane (Schleicher & Schuell NA45)
0.025-μm filter discs (Millipore VS)

Additional reagents and equipment for mobility-shift DNA binding assay (see Support Protocol)

Prepare biotinylated, labeled DNA fragment

1. Digest 50 μg of plasmid DNA in 100 μl with one or more appropriate restriction endonucleases to obtain DNA probe.
2. Add the following to the probe mixture:

   Biotin-11-dUTP to a final concentration of 20 μM
   Radioactive dNTP for incorporation into the 5′ overhang
   100-fold molar excess of corresponding unlabeled dNTP
   Remaining two unlabeled dNTPs to a final concentration of 200 μM
   5 U Klenow fragment.

   This reaction is identical to that used in standard DNA-binding assays (e.g., see Support Protocol), except biotinylated dUTP is incorporated into one end of the DNA fragment in place of TTP and the fragment is radiolabeled to low instead of high specific activity.

3. Precipitate the biotinylated probe. Isolate the probe by agarose gel electrophoresis using DEAE membrane.

   Gel purification of the biotinylated fragment removes unreacted biotin-11-dUTP. This is essential because any free biotin-dUTP binds to streptavidin and then to the biotin-cellulose column, reducing the apparent capacity of both materials to react with protein.

4. Resuspend the probe in TE buffer and measure an aliquot for Cerenkov counts. Estimate the DNA concentration by ethidium bromide dot quantitation.
5. Test the biotinylated probe to be certain it will efficiently bind to the protein of interest. Use a standard binding assay (e.g., see Support Protocol) with the biotinylated fragment as probe.

Prepare biotin-cellulose resin

6. Place 200 μl biotin-cellulose in a 1.5-ml microcentrifuge tube. Spin the resin for 30 sec in a microcentrifuge and remove the supernatant. Add to the pellet:

   1.0 ml biotin-cellulose binding buffer
   500 μg/ml BSA
   200 μg carrier DNA

CU IASI CENTRAL UNIVERSITY LIBRARY

Gently mix the tube 5 min on a rotating wheel.

This step blocks nonspecific protein and nucleic acid binding sites present on the biotin-cellulose resin.

7. Spin the resin, remove the supernatant, and resuspend pellet in 1.0 ml of biotin-cellulose elution buffer. Gently mix 5 min on a rotating wheel. Repeat this wash.

   Washing removes molecules on the biotin-cellulose resin that might later be eluted from the resin by the biotin-cellulose elution buffer.

8. Spin the resin, remove the supernatant, and resuspend pellet in 1.0 ml of biotin-cellulose binding buffer. Repeat this wash.

   This 1:6 dilution of pretreated biotin-cellulose is ready for use and can be stored for several months.

Set up binding reaction

9. Determine the molar concentration of the protein to be purified via a mobility-shift assay (e.g., see Support Protocol).

10. Set up a standard binding reaction containing the protein to be purified, carrier DNA, and a 10-fold molar excess of biotinylated fragment relative to the protein to be purified. Allow the reaction to go to completion for ~15 min.

    Use reaction conditions that optimize protein binding to its recognition site. The composition of the biotin-cellulose binding buffer should be the same as that optimized for protein-DNA binding.

11. Add a 5-fold molar excess of streptavidin relative to the biotinylated fragment. Continue the binding reaction for an additional 5 min at 30°C.

Bind protein/DNA/streptavidin complex to biotin-cellulose resin

12. In a separate tube, place 2 µl pretreated biotin-cellulose (12 µl of the 1:6 dilution) for each picomole of biotinylated DNA fragment in the binding reaction. Spin the resin and remove the supernatant.

    One or two microliters of biotin-cellulose can easily be seen at the bottom of the microcentrifuge tube.

13. Transfer the binding reaction mix into the tube with the biotin-cellulose resin using a pipettor. Gently resuspend the resin and incubate on a rotating wheel for 30 min.

    This incubation can be done either at 4°C or at room temperature, depending on the stability of the protein.

14. Spin the resin and remove the supernatant.

    The supernatant should be measured using a mobility shift assay (e.g., see Support Protocol) to determine what percentage of the biotinylated fragment and the protein has been removed from the supernatant. It is also useful to assay the supernatant from the binding reaction for a control DNA-binding protein to determine whether the protein of interest has been specifically depleted or whether multiple DNA-binding proteins have been depleted nonspecifically. The latter observation would suggest that the matrix is acting as a nonspecific DNA-affinity column. For some applications, a protein fraction specifically depleted for a particular DNA-binding protein is a valuable reagent.

Wash the resin

15. Resuspend the biotin-cellulose pellet in 500 µl biotin-cellulose binding buffer. Mix by gently inverting the tube 1 to 2 min. Spin the resin and remove the supernatant.

BU AMI CENTRAL UNIVERSITY LIBRARY

Repeat this procedure twice. The second time, transfer to a clean microcentrifuge tube.

Transferring the reaction avoids elution of proteins that were bound nonspecifically to the walls of the tube in the first binding incubation.

Elute the protein

16. Resuspend the biotin-cellulose pellet in at least an equal volume of biotin-cellulose elution buffer. Mix gently on a rotating wheel for 20 min.

    The salt concentration in the biotin-cellulose elution buffer must be determined empirically. Successively higher salt concentrations may be tested until the concentration of eluted protein is maximal.

17. Spin the resin. Save the supernatant and assay for binding activity.

    Small volumes of protein solutions can be dialyzed effectively on 0.025-μm filter discs.

ALTERNATE PROTOCOL 1

PURIFICATION USING A MICROCOLUMN

Although the batch method in the basic protocol is rapid and well-suited for analytical-scale purification, larger volumes of biotin-cellulose resin can be better handled in a microcolumn. This method is also used to elute the protein in as small a volume (i.e., as high a concentration) as possible.

***Additional Materials** (also see Basic Protocol)*

Silanized glass wool
1.0-ml pipet tip
Ring stand

1. Prepare the biotinylated DNA fragment and biotin-cellulose resin and set up binding reaction (see Basic Protocol, steps 1 to 11).

2. Place a small plug of silanized glass wool in the bottom of a 1.0-ml pipet tip. Firmly attach the pipet-tip microcolumn to a ring stand.

   Prewet the glass wool in biotin-cellulose binding buffer before insertion into the pipet tip to avoid trapping air bubbles which might denature proteins.

3. Add 500 μl binding buffer to the microcolumn. Maintain a steady flow through the glass wool plug.

   If the column does not flow smoothly, a pipettor can be gently inserted into the top of the microcolumn. Slightly depressing the plunger will start the column or increase the flow.

4. Add at least 40 μl of 1:1 biotin-cellulose slurry to the microcolumn. Allow the buffer to run down to the surface of the resin.

5. Equilibrate the resin with 3 column volumes of biotin-cellulose binding buffer if the resin has already been pretreated (Basic Protocol, steps 6 to 8). If the resin has not been pretreated, wash sequentially with 3 column volumes each of biotin-cellulose binding buffer, biotin-cellulose binding buffer with 500 μg/ml BSA and 200 μg/ml poly(dI-dC)·poly(dI-dC), and biotin-cellulose elution buffer. Finally, equilibrate with biotin-cellulose binding buffer.

   The biotin-cellulose microcolumn can be washed very rapidly. Each wash takes 2 or 3 min. The drop size from the pipet tip is ~25 μl but this will change with alterations in the ionic strength and protein concentration of the eluate.

CU IAS/CENTRAL UNIVERSITY LIBRARY

6. Load the binding reaction mix (Basic Protocol, step 11) onto the microcolumn and collect the flowthrough.

   The column can be run as fast as 6 to 10 column volumes/hr without affecting the amount of biotinylated fragment bound by the resin. If the flow rate is too slow, use a pipettor to apply pressure to the column. If the flow rate is too fast, plug the tip of the microcolumn with Parafilm in between drops.

7. Wash with 4 column volumes of biotin-cellulose binding buffer. Discard the flow-through.

8. Wash with 3 column volumes of biotin-cellulose elution buffer. Collect 2-drop fractions and assay.

   Fractions as small as 5 µl may be rapidly and effectively dialyzed on 0.025-µm filter discs with minimal loss of volume and activity. Float a filter (shiny side up) in a petri dish on top of 20 ml dialysis buffer. Allow the filter 10 min to wet. Place the sample (5 to 100 µl) to be dialyzed onto the surface of the filter. Surface tension will keep the sample confined in a drop unless there is a detergent in the sample. After 1 hr remove the sample. Once the protein has been eluted from the resin, the resin is effectively the same as a sequence-specific DNA-affinity microcolumn.

PURIFICATION USING STREPTAVIDIN-AGAROSE

ALTERNATE PROTOCOL 2

When high-quality free streptavidin is not available or cellulose is an inappropriate resin, a simple variation on the basic protocol may be employed. In this protocol, the same biotinylated DNA fragment is used but is removed from solution directly by streptavidin-agarose (see Fig. 9.7.2).

Additional Materials (also see Basic Protocol)

Streptavidin-agarose

1. Prepare biotinylated DNA fragment and resin, substituting streptavidin-agarose for biotin-cellulose, and set up binding reaction (see Basic Protocol, steps 1 to 11).

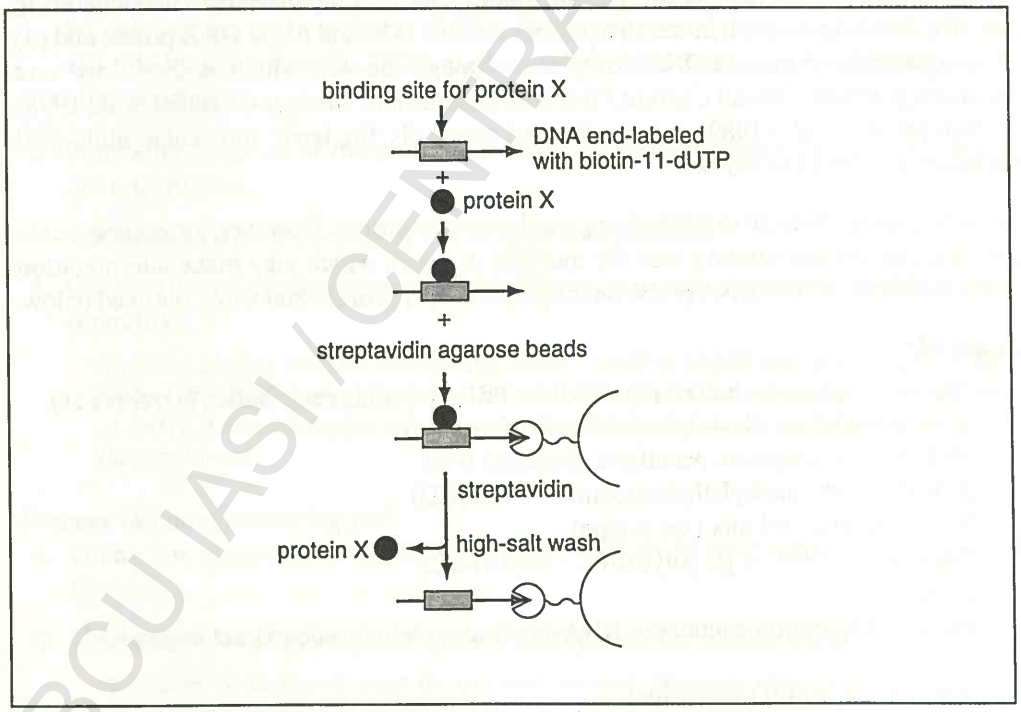


Figure 9.7.2 Purification of DNA-binding proteins using streptavidin-agarose.

RCU IASI / CENTRAL UNIVERSITY LIBRARY

2. In a separate tube, add 50 µl pretreated streptavidin-agarose (300 µl of the 1:6 dilution) for each picomole of biotinylated DNA fragment in the binding reaction. Spin the resin and remove the supernatant.
3. Transfer the binding-reaction mix into the tube with the streptavidin-agarose using a pipettor. Gently resuspend the resin and incubate on a rotating wheel for 30 min to 2 hr.
4. Wash and elute the protein (see Basic Protocol, steps 14 to 17).

Like biotin-cellulose, streptavidin-agarose may also be used in a microcolumn. Follow the microcolumn alternate protocol.

SUPPORT PROTOCOL

MOBILITY-SHIFT ASSAY

The DNA-binding assay using nondenaturing polyacrylamide gel electrophoresis (PAGE) provides a simple, rapid, and extremely sensitive method for detecting sequence-specific DNA-binding proteins. Proteins that bind specifically to an end-labeled DNA fragment retard the mobility of the fragment during electrophoresis, resulting in discrete bands corresponding to the individual protein-DNA complexes. The assay can be used to test binding of purified proteins (see Basic Protocol) or of uncharacterized factors found in crude extracts. This assay also permits quantitative determination of the affinity, abundance, association rate constants, dissociation rate constants, and binding specificity of DNA-binding proteins.

The utility of this technique is underscored by the many proteins that have been characterized using this assay. It has become clear that there is no single protocol that works best for all proteins. Rather, several variables can be changed to optimize binding. There are several options available for the design of the DNA probe, the binding reaction conditions, and the gel running conditions. The reader is referred to Buratowski and Chodosh (1996) for additional details and protocols.

This protocol can be divided into four stages: (1) preparation of a radioactively labeled DNA probe containing a particular protein binding site; (2) preparation of a nondenaturing gel; (3) a binding reaction in which a protein mixture is bound to the DNA probe; and (4) electrophoresis of protein-DNA complexes through the gel, which is then dried and autoradiographed. Consult *Current Protocols in Molecular Biology* (Ausubel et al., 1998) or Sambrook et al. (1989) for details and protocols for basic molecular biological techniques referred to below.

DNA fragments from 20 to 300 bp long may be used as probes. However, longer fragments are likely to contain binding sites for multiple proteins, which may make interpretation of the gel difficult. The DNA probe can be prepared in one of several ways, outlined below.

Materials

10× electrophoresis buffer: e.g., TAE or TBE electrophoresis buffer (*APPENDIX 2E*) or Tris-glycine electrophoresis buffer (*UNIT 10.3*)
30% (w/v) ammonium persulfate, prepared fresh
N,N,N′,N′-tetramethylethylenediamine (TEMED)
Nondenaturing gel mix (see recipe)
Bulk carrier DNA, e.g., poly(dI-dC)·poly(dI-dC)
BSA
Protein preparation containing DNA-binding protein (crude extract or purified fraction)
10× loading buffer (see recipe)

PSU IAST / PENNSTATE UNIVERSITY LIBRARY

Constant-temperature water bath
Two-head peristaltic pump
10-µl glass capillary pipet (optional)
Clay-Adams screw-top loader (optional)
Whatman 3MM filter paper (or equivalent)

Additional reagents and equipment for digesting DNA with restriction endonucleases, DNA labeling, agarose and nondenaturing polyacrylamide gel electrophoresis, recovery of DNA from gels, oligonucleotide synthesis, PCR, ethanol precipitation, ethidium bromide dot quantitation, and autoradiography (see Ausubel et al., 1998, or Sambrook et al., 1989)

Prepare the DNA probe

1a. *For restriction endonuclease fragments:* Isolate a small DNA fragment containing the binding site of interest from a plasmid using a standard restriction endonuclease digestion. Label the fragment by filling in a 5′ overhang with the Klenow fragment of *Escherichia coli* DNA polymerase and ^{32}P-labeled nucleotide or by end labeling using polynucleotide kinase. Separate the fragment from the plasmid by gel electrophoresis.

Kinased probes should be avoided in experiments using crude protein preparations that might contain phosphatase activity.

Agarose gels are useful for resolving fragments as small as 50 bp. Fragments can be recovered using low-melting-temperature agarose or DEAE paper. For smaller DNA fragments, nondenaturing polyacrylamide gels can be used.

1b. *For synthetic oligonucleotides:* Synthesize and anneal complementary oligonucleotides to generate a double-stranded DNA fragment containing the binding site of interest. Label the probe using polynucleotide kinase.

The kinase reaction can be performed before annealing to label only one strand, or after annealing to label both strands. Alternatively, the oligonucleotides can be designed to leave an overhang that can be filled in with labeled nucleotide and Klenow fragment.

1c. *For PCR fragments:* Generate a DNA fragment containing the binding site of interest by polymerase chain reaction (PCR). End-label one of the primers with polynucleotide kinase before the PCR reaction, or label the double-stranded PCR product after purification.

2. Following isolation of the probe, determine its concentration by ethidium bromide dot quantitation.

Concentrations in the range of 2 to 50 ng/µl are convenient.

3. Count 1 µl for Cerenkov counts in a scintillation counter to determine specific activity (cpm/µl).

A typical binding reaction will contain about ~5,000 to 20,000 cpm and ~10 to 100 fmol probe (10 fmol DNA in a final reaction volume of 10 µl gives a total DNA concentration of 1 nM). If desired, probes can be stored up to 4 to 6 weeks at 4°C before proceeding with the experiment.

Prepare the nondenaturing gel

4. Dilute 10× electrophoresis buffer to prepare enough 1× electrophoresis buffer to fill the tank.

5. Assemble washed glass plates and 1.5-mm spacers for casting the gel.

All traces of detergent must be removed because detergent will disrupt protein-DNA interactions.

BOSTON UNIVERSITY MEDICAL CENTER LIBRARY

6. Add 150 µl of 30% ammonium persulfate and 70 µl TEMED to 60 ml nondenaturing gel mix prepared using the same buffer as that used for electrophoresis. Swirl gently to mix.

   The amounts of ammonium persulfate, TEMED, and nondenaturing gel mix can be scaled up or down as necessary depending on the size and number of gels.

7. Pour the gel mix between the plates and insert a comb. Allow the gel to completely polymerize for 20 min.

   For optimal results, use a comb with teeth that are ≥7 mm wide.

8. Remove the comb and bottom spacer and attach the plates to the electrophoresis tank after filling the lower reservoir with 1× electrophoresis buffer. Fill the upper reservoir of the tank with 1× electrophoresis buffer. With a bent-needle syringe, remove any air bubbles trapped beneath the gel and flush out the wells. Prerun the gel 30 to 60 min at 100 V.

   For low-ionic-strength buffers (≤0.5×), use a pump with two heads and a flow rate of 5 to 30 ml per min to exchange buffer between the upper and lower reservoirs. Recirculation of the buffer is essential to prevent polarization due to the low buffering capacity of the buffer.

Prepare the binding reactions

9. While the gel is prerunning, assemble the binding reaction by combining the following in a 0.5-ml or 1.5-ml microcentrifuge tube:

   5,000 to 20,000 cpm radiolabeled probe DNA (0.1 to 0.5 ng, ≥10 fmol)
   0.1 to 2 µg nonspecific carrier DNA
   300 µg/ml BSA
   ≥10% (v/v) glycerol
   Appropriate buffer and salt
   Water or buffer to obtain a final reaction volume of 10 to 15 µl
   DNA-binding protein (~15 µg crude extract or ~5 to 25 ng purified protein).

   The protein should be added last.

10. Mix gently by tapping the bottom of the tube with a finger.

    Avoid introducing bubbles in the mix.

11. Incubate the binding reaction mix 15 to 30 minutes in a constant-temperature water bath.

    Optimal incubation temperatures for different proteins can vary from room temperature to 37°C.

Run the gel

12. Load each binding reaction into the appropriate well of the prerun gel using either a 10-µl glass capillary pipet and Clay-Adams screw-top loader, or a pipettor. Load a small volume of 10× loading buffer with dyes into a separate well.

    The dyes are used to monitor the progress of electrophoresis.

    There is no stacking gel in this system, so precise loading with little mixing with the gel buffer is necessary to obtain sharp bands on the gel. Allow the sample to fall along one side of the well to prevent dilution and avoid bubbles in the well.

13. Electrophorese at ~30 to 35 mA for the minimum time required to give good separation of free probe and the protein-DNA complexes. Stop the gel before the bromphenol blue approaches the bottom of the gel (~1.5 to 2 hr for a 15- to 20-cm gel).

    Longer run times may cause a weaker signal due to partial dissociation of complexes during electrophoresis.

CSU IASL CENTRAL UNIVERSITY LIBRARY

Bromphenol blue migrates at approximately the same position as a 70-bp DNA probe. For probes <70 bp, do not run the bromphenol blue to the bottom of the gel.

If electrophoresis is performed at room temperature, the glass plates should be allowed to become only slightly warm. Decrease the voltage if the plates become any hotter. To run the gel faster, put the apparatus in a cold room, so that higher voltages may be used without heating the glass plates.

Analyze the gel

14. Remove the glass plates from the gel box and carefully remove the side spacers.
15. Using a spatula, slowly pry the glass plates apart, allowing air to enter between the gel and the glass plate.

    The gel should remain attached to only one of the plates. Prying the plates apart too quickly may tear the gel or cause it to stick to both plates. If this occurs or if the gel has become distorted, squirt a stream of water underneath it. This will reduce the stickiness of the gel. Be careful not to let the gel slide off the plate.

16. Lay the glass plate with the gel attached on the bench with the gel facing up. Place three sheets of Whatman 3MM filter paper cut to size on top of the gel.
17. Support both sides of the gel sandwich and carefully flip it over so that the filter paper is on the bottom and the glass plate is on the top.
18. Carefully lift up one end of the glass plate. Peel the filter paper with the gel attached to it from the plate.
19. Cover the gel with plastic wrap and dry under vacuum with heat.
20. Autoradiograph the dried filter overnight without an intensifying screen or 2 to 3 hr with an intensifying screen (*UNIT 10.11*).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Biotin-cellulose binding buffer

12% glycerol
12 mM HEPES-NaOH, pH 7.9
4 mM Tris·Cl, pH 7.9 (*APPENDIX 2E*)
60 mM KCl
1 mM EDTA
1 mM DTT
Store at –20°C upon addition of DTT

This is a typical binding buffer. The composition of binding buffer—especially with respect to pH, ionic strength, and the presence or absence of $MgCl_2$—should be determined by those conditions which optimize the binding of the protein of interest to its recognition site.

Biotin-cellulose elution buffer

12% glycerol
20 mM Tris·Cl, pH 6.8 (*APPENDIX 2E*)
1 M KCl
5 mM $MgCl_2$
1 mM EDTA
1 mM DTT
200 µg/ml BSA
Store at –20°C upon addition of DTT

continued

B.C.U. „M. Eminescu" IAŞI
CHIMIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

This is a typical elution buffer. The composition of elution buffer—especially with respect to pH, ionic strength, and the presence or absence of $MgCl_2$—should be determined by those conditions that maximize the dissociation rate of the protein from its recognition site. Another carrier protein, such as insulin or hemoglobin, may be substituted for BSA. The small size of insulin can be useful if protein gels will be run to determine the size of the regulatory factor.

Loading buffer, 10×

20% Ficoll 400
0.1 M Na_2EDTA, pH 8 (*APPENDIX 2E*)
1.0% sodium dodecyl sulfate (SDS; *APPENDIX 2E*)
0.25% bromphenol blue
0.25% xylene cyanol (optional)

Xylene cyanol runs ~50% as fast as bromphenol blue and can interfere with visualization of bands of moderate molecular weight, but can be helpful for monitoring very long runs)

Nondenaturing gel mix

For 60 ml:
6.0 ml 10× TAE or TBE elecrophoresis buffer (*APPENDIX 2E*) or Tris-glycine electrophoresis buffer (*UNIT 10.3*)
8.1 ml 30% (w/v) acrylamide
3.0 ml 2% (w/v) bisacrylamide
42.9 ml H_2O
Prepare fresh for each use

CAUTION: *Acrylamide monomer is neurotoxic. Gloves should be worn when handling the solution, and the solution should not be pipetted by mouth.*

COMMENTARY

Background Information

The biotin/streptavidin purification method works well because the interaction between biotin and avidin (or avidin-like proteins) is one of the strongest known noncovalent interactions. The dissociation constant for the streptavidin-biotin complex is ~10^{-15} M. Avidin, from egg white, and streptavidin, from *Streptomyces avidinii*, are tetrameric proteins containing four high-affinity binding sites for the vitamin biotin. Since streptavidin is multivalent, it is able to serve as a bridge between the biotinylated DNA fragment and the biotin-containing resin. The strong interaction is extremely useful for purification of DNA-binding proteins, because DNA-affinity columns with streptavidin/biotin bridges can be washed under a wide variety of conditions (i.e., 2 M KCl and 1% SDS) without removing either the streptavidin or the biotinylated DNA fragment from the matrix.

Two properties of streptavidin make it more suitable than avidin for use in DNA-affinity purification. The first is that streptavidin, unlike avidin, is not a glycoprotein, and the second is that streptavidin is slightly acidic whereas avidin is basic. Therefore, streptavidin is less likely to bind nonspecifically to cellular glycoproteins and to acidically charged cell components such as nucleic acids.

Several factors make this method simple, rapid, and effective. The same binding conditions and DNA fragment used in the mobility-shift DNA-binding assay to identify a protein (see Support Protocol; Buratowski and Chodosh, 1996) can be used to effect its purification. In addition, the binding of the protein to its DNA recognition site in solution is more efficient than protein-DNA interactions that take place on a column matrix (discussed below). Most importantly, binding in solution allows each reaction parameter to be optimized on an analytical scale by using the gel binding assay.

As an analytical technique, biotin/streptavidin DNA-affinity purification permits the direct identification of a wide variety of sequence-specific DNA-binding proteins. It has already been successfully used to identify hormone receptors, components in mRNA splicing complexes, and RNA polymerase II and III transcription factors (Haeuptle et al., 1983; Chodosh et al., 1986; Grabowski and Sharp, 1986; Kasher et al., 1986).

CSU IAS CENTRAL UNIVERSITY LIBRARY

Another method is commonly used for purifying DNA-binding proteins, whereby DNA-protein binding interactions occur on a column matrix. In this technique, catenated DNA-binding sites are covalently coupled with cyanogen bromide to Sepharose CL-2B (Kadonaga and Tjian, 1986; Rosenfeld and Kelly, 1986). A partially purified protein fraction is combined with competitor DNA and passed through a DNA-Sepharose resin, binding the protein to the surface of the matrix. Generally, multiple passes through the column are used to effect purification, with 25- to 50-fold purification achieved with each column passage. The biotin/streptavidin technique, on the other hand, typically yields 50- to 250-fold purification.

Critical Parameters

Only one end of the DNA fragment should be labeled with biotin. If both ends are labeled, the biotinylated fragment may bind to the resin in such a fashion as to sterically displace the protein bound to the DNA fragment. Thus, one of the restriction enzymes used must be chosen to permit the incorporation of biotin-11-dUTP in place of TTP. Furthermore, since only one end of the fragment is biotin-labeled, the restriction enzyme generating the nonbiotinylated end should leave no overhang, a 3′ overhang, or a 5′ overhang that will not incorporate biotin-dUTP. For binding sites cloned into the pUC polylinker, an *Eco*RI and *Hinc*II fragment fulfills these requirements. Biotin-11-dUTP may be incorporated into the *Eco*RI overhang along with a low-specific-activity dATP radioactive label. The blunt *Hinc*II end of the fragment remains unlabeled.

Alternately, if both ends have 5′ overhangs which permit incorporation of biotin-dUTP, the plasmid may first be digested with one of the enzymes and end-labeled with biotin-dUTP and a radioactive dNTP using the Klenow fragment. After heat-inactivating the Klenow fragment, the plasmid may then be digested with the second restriction endonuclease. Use of annealed complimentary oligonucleotides with the desired ends obviates the need to find appropriate restriction sites.

The low-specific-activity label incorporated into the biotinylated DNA fragment is used to help monitor both the binding of the protein to the biotinylated fragment and the binding of the biotinylated fragment to the biotin-cellulose resin via a streptavidin bridge. The amounts of the radioactive dNTP and the nonradioactive dNTP added to the Klenow reaction should be in a ratio such that 1% to 2% of the biotinylated fragments are labeled. For example, 100 μCi of 6000 Ci/mmol dATP (0.2 μM) may be added to a reaction containing 20 μM unlabeled dATP.

If the radioactive dNTP and the biotin-dUTP residue are incorporated on the same end of the DNA fragment, as in an *Eco*RI site, decay of the radioactive nuclide will result in formation of an unlabeled, nonbiotinylated probe that binds to the protein but is incapable of binding to streptavidin. This molecule is a competitive inhibitor during protein purification and can be kept to a minimum by labeling to a lower specific activity.

Over time, streptavidin tetramers dissociate into monomers. The integrity of the tetramer may be monitored by performing native (nondenaturing) gel electrophoresis. Since streptavidin monomers will bind to the biotinylated DNA fragment, but will be unable to bind to the biotin-cellulose resin, the monomeric form of streptavidin will inhibit purification.

Most proteins can be eluted from their binding site with 1 M KCl. Since the added elution buffer is diluted by the binding buffer in the biotin-cellulose pellet, the actual salt concentration in the resuspended pellet will be lower than that of the elution buffer. As a result, it is useful to remove a small aliquot, dilute it, and measure the actual salt concentration with a conductivity meter. Otherwise, the time, temperature, and composition of the buffer used in the elution step should be chosen to maximize the protein's dissociation rate from its binding site. Thus, if a protein's dissociation rate is 15 sec in 1 M KCl at room temperature, and 3 hr in 1 M KCl at 4°C, a 5-min incubation in 1 M KCl at room temperature should be sufficient to elute the protein. Dissociation rates can easily be measured using a mobility shift DNA-binding assay described in this chapter (Support Protocol; Buratowski and Chodosh, 1996).

Troubleshooting

A major advantage of the biotin/streptavidin method is the ability to monitor the purification on an analytical scale using a mobility-shift assay. Because different parameters can be independently examined, it is possible to pinpoint the particular step that is not working. Several parameters can be examined routinely. Using the biotinylated radiolabeled fragment, the specific binding of the DNA fragment to its protein can be directly determined. In addition, the binding conditions can be optimized in a binding reaction containing the DNA probe and protein. The probe can be a biotinylated or a

BOSTON BIOMEDICAL RESEARCH INSTITUTE LIBRARY

nonbiotinylated radiolabeled DNA fragment. The complex is then resolved by native (nondenaturing) gel electrophoresis.

Streptavidin binding to the probe can be monitored and titrated by varying the amount of streptavidin in a binding reaction. The resulting streptavidin-DNA complex is resolved from free DNA by native gel electrophoresis.

Binding of this complex to the biotin-cellulose matrix can be directly monitored. First, biotin-cellulose is incubated with the complex. Next the resin is spun and aliquots of the supernatant are assayed for the presence of the radiolabeled fragment on a polyacrylamide gel. This basic procedure allows several relevant parameters to be examined, including the amount of streptavidin required for complete binding of the fragment to the resin, the amount of biotin-cellulose required to bind a given amount of fragment, the binding time, and the optimal binding conditions. These three final steps can also be monitored for the specific adsorption of the protein to the column. By assaying the amount of binding activity in the supernatant before and after addition of biotin cellulose, the adsorption efficiency can be determined.

Nonspecific resin binding can be examined by assaying the binding activity in a supernatant containing the protein, the DNA carrier poly(dI-dC)·poly(dI-dC), streptavidin, and biotin-cellulose, but lacking the biotinylated DNA fragment with the protein's binding site. Alternatively, a nonspecific biotinylated DNA fragment can be added. Interestingly, nonspecific protein binding to the resin is minimized with a lower poly(dI-dC)·poly(dI-dC) concentration than is required to generate a discrete band in the mobility shift assay.

The conditions for protein elution from the column can also be directly examined and optimized. After protein adsorption onto the matrix, elution parameters such as time, temperature, pH, carrier protein, and ionic strength can be varied. The amount of protein released into the supernatant is then assayed by the mobility shift assay. In optimizing this step, it is very helpful to know as much as possible about the variables affecting the dissociation rate of the protein from its DNA binding site. These can be studied quickly and rigorously with the mobility shift assay. Troubleshooting guidelines for this assay are presented in Buratowski and Chodosh (1996).

Anticipated Results

The strong interaction between biotin and avidin can result in 25 nmol of biotinylated DNA fragment binding to 1 ml of biotin-cellulose resin. This corresponds to 1.5 mg of a 60 kDa polypeptide. After parameter optimization, ~50- to 250-fold purification of the DNA-protein complex can be obtained. Once the protein has been adsorbed to the column >99% of the remaining protein can be removed with repeated column washing using the binding buffer. Moreover, ~30% to 80% of the binding activity can be recovered from the column. The recovery efficiency depends on how effectively nonspecific binding sites have been blocked. This purification scheme may be used either with crude extracts or with partially purified protein fractions.

Time Considerations

Starting with a biotinylated DNA fragment, the batch protocol (Basic Protocol) can be performed in ~1.5 hr, the microcolumn (Alternate Protocol 1) in ~2.5 hr, and the streptavidin-agarose (Alternate Protocol 2) protocol in ~1.5 hr. The DNA-binding assay (Support Protocol) may take 4 to 5 hr to complete from the time the gel is poured to the time it is dried onto filter paper. Only a small fraction of this time is labor intensive. Thus, several DNA-binding gels can be run simultaneously while other experiments are being performed. Furthermore, the results are often available within 2 to 3 hr after completion of the experiment.

Literature Cited

Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) 1998. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Buratowski, S. and Chodosh, L. 1996. Mobility-shift DNA-binding assay using gel electrophoresis. *In* Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 12.2.1-12.2.11. John Wiley & Sons, New York.

Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late promoter. *Mol. Cell. Biol.* 6:4723-4733.

Grabowski, P.J. and Sharp, P.A. 1986. Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science* 233:1294-1299.

Haeuptle, M.-T., Aubert, M.L., Kjiane, J., and Kraehenbuhl, J.-P. 1983. Binding sites for lactogenic and somatogenic hormones from rabbit mammary gland and liver. *J. Biol. Chem.* 258:305-314.

Kadonaga, J.T. and Tjian, R. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 83:5889-5893.

Kasher, M.S., Pintel, D., and Ward, D.C. 1986. Rapid enrichment of HeLa transcription factors IIIB and IIIC by using affinity chromatography based on avidin-biotin interactions. *Mol. Cell. Biol.* 6:3117-3127.

Rosenfeld, P.J. and Kelly, T.J. 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261:1398-1408.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Key References

Chodosh et al., 1986. See above.

Describes the biotin/streptavidin-agarose microcolumn procedure from which the second alternate protocol of this unit was drawn.

Kasher et al., 1986. See above.

Describes a biotin-cellulose/streptavidin procedure similar to the basic protocol of this unit.

Contributed by Lewis A. Chodosh
University of Pennsylvania
Philadelphia, Pennsylvania

Stephen Buratowski
Harvard Medical School
Boston, Massachusetts

BCU IASI / CENTRAL UNIVERSITY LIBRARY

ELECTROPHORESIS

10

INTRODUCTION 10.0.1

10.1 One-Dimensional SDS Gel Electrophoresis of Proteins 10.1.1
Electricity and Electrophoresis 10.1.1
Basic Protocol 1: Denaturing (SDS) Discontinuous Gel Electrophoresis: Laemmli Gel Method 10.1.3
Alternate Protocol 1: Electrophoresis in Tris-Tricine Buffer Systems 10.1.9
Alternate Protocol 2: Nonurea Peptide Separations with Tris Buffers 10.1.11
Alternate Protocol 3: Continuous SDS-PAGE 10.1.13
Alternate Protocol 4: Casting and Running Ultrathin Gels 10.1.14
Support Protocol 1: Casting Multiple Single-Concentration Gels 10.1.15
Alternate Protocol 5: Separation of Proteins on Gradient Gels 10.1.17
Support Protocol 2: Casting Multiple Gradient Gels 10.1.21
Basic Protocol 2: Electrophoresis in Single-Concentration Minigels 10.1.23
Support Protocol 3: Preparing Multiple Gradient Minigels 10.1.26
Reagents and Solutions 10.1.27
Commentary 10.1.29

10.2 One-Dimensional Isoelectric Focusing of Proteins in Slab Gels 10.2.1
Basic Protocol: Isoelectric Focusing in Slab Gels Under Denaturing Conditions 10.2.1
Support Protocol 1: Electroblotting from Denaturing Isoelectric-Focusing Slab Gels 10.2.4
Support Protocol 2: Fluorography of Denaturing Isoelectric-Focusing Slab Gels 10.2.5
Reagents and Solutions 10.2.6
Commentary 10.2.6

10.3 One-Dimensional Electrophoresis Using Nondenaturing Conditions 10.3.1
Basic Protocol: Continuous Electrophoresis in Nondenaturing Polyacrylamide Gels 10.3.1
Alternate Protocol: Native Discontinuous Electrophoresis and Generation of Molecular Weight Standard Curves (Ferguson Plots) 10.3.5
Reagents and Solutions 10.3.8
Commentary 10.3.9

10.4 Two-Dimensional Gel Electrophoresis 10.4.1
Basic Protocol 1: High-Resolution Equilibrium Isoelectrofocusing in Tube Gels 10.4.2
Support Protocol 1: Conducting pH Profile Measurements 10.4.7
Alternate Protocol 1: Nonequilibrium Isoelectrofocusing of Very Acidic Proteins 10.4.8
Alternate Protocol 2: Nonequilibrium Isoelectrofocusing of Basic Proteins 10.4.8
Basic Protocol 2: Isoelectrofocusing Using Immobilized pH Gradient Gel Strips 10.4.9
Support Protocol 2: Electrophoresis on Immobilized pH Gradient Gels 10.4.13
Support Protocol 3: Casting an Immobiline Gel 10.4.16
Support Protocol 4: Preparing Tissue Culture Cell Extracts for Isoelectrofocusing 10.4.17
Support Protocol 5: Preparing Proteins in Tissue Samples 10.4.19

continued

BCU IAHT CENTRAL UNIVERSITY LIBRARY

Copyright © 1998 by John Wiley & Sons, Inc.

Basic Protocol 3: Second-Dimension Electrophoresis of IEF Tube Gels 10.4.20
Basic Protocol 4: Second-Dimension Electrophoresis of IPG Gels 10.4.22
Support Protocol 6: Preparing Molecular Weight Standards for Two-Dimensional Gels 10.4.24
Alternate Protocol 3: Diagonal Gel Electrophoresis (Nonreducing/Reducing Gels) 10.4.25
Support Protocol 7: Using Two-Dimensional Protein Databases 10.4.26
Reagents and Solutions 10.4.27
Commentary 10.4.30

10.5 Protein Detection in Gels Using Fixation 10.5.1
Basic Protocol 1: Coomassie Blue R-250 Staining 10.5.1
Alternate Protocol 1: Rapid Coomassie Blue G-250 Staining 10.5.2
Alternate Protocol 2: Acid-Based Coomassie Blue G-250 Staining 10.5.3
Basic Protocol 2: Silver Staining 10.5.4
Alternate Protocol 3: Nonammoniacal Silver Staining 10.5.5
Alternate Protocol 4: Rapid Silver Staining 10.5.6
Alternate Protocol 5: Enhanced-Background (Two-Stage) Rapid Silver Staining 10.5.7
Support Protocol: Gel Photography 10.5.8
Reagents and Solutions 10.5.9
Commentary 10.5.11

10.6 Protein Detection in Gels Without Fixation 10.6.1
Basic Protocol 1: Protein Detection in Gels by SDS Precipitation 10.6.1
Basic Protocol 2: Preparation of Contact Blots 10.6.2
Basic Protocol 3: IAEDANS Modification of Proteins 10.6.3
Basic Protocol 4: Fluorescamine Labeling of Proteins 10.6.5
Reagents and Solutions 10.6.6
Commentary 10.6.6

10.7 Electroblotting from Polyacrylamide Gels 10.7.1
Basic Protocol 1: Electroblotting onto PVDF Membranes 10.7.1
Alternate Protocol 1: Electroblotting of Proteins for Sequence Analysis 10.7.5
Alternate Protocol 2: Electroblotting onto Nitrocellulose Membranes 10.7.6
Alternate Protocol 3: Protein Electroblotting in Semidry Systems 10.7.6
Basic Protocol 2: Protein Elution from PVDF Membranes Using Detergents 10.7.8
Alternate Protocol 4: Protein Elution from PVDF Membranes Using Acidic Extraction with Organic Solvents 10.7.10
Reagents and Solutions 10.7.11
Commentary 10.7.11

10.8 Detection of Proteins on Blot Membranes 10.8.1
Basic Protocol 1: Amido Black Staining 10.8.2
Basic Protocol 2: Coomassie Blue R-250 Staining 10.8.2
Basic Protocol 3: Ponceau S Staining 10.8.3
Basic Protocol 4: Colloidal Gold Staining 10.8.3
Basic Protocol 5: Colloidal Silver Staining 10.8.4
Basic Protocol 6: India Ink Staining 10.8.4
Basic Protocol 7: Fluorescamine Labeling 10.8.5
Alternate Protocol: IAEDANS Labeling 10.8.5
Commentary 10.8.6

continued

BCU IASI / CENTRAL UNIVERSITY LIBRARY

10.9 Capillary Electrophoresis of Proteins and Peptides 10.9.1
Instrumentation 10.9.2
Separation Theory 10.9.3
Strategic Planning 10.9.4
Basic Protocol 1: Separation of Proteins by Isoelectric Focusing 10.9.5
Basic Protocol 2: Separation of Proteins 10.9.6
Basic Protocol 3: Analytical Peptide Separations 10.9.8
Basic Protocol 4: Micropreparative Capillary Electrophoresis: Multiple Separations 10.9.9
Alternate Protocol: Micropreparative Capillary Electrophoresis: Single Separation 10.9.10
Commentary 10.9.11

10.10 Immunoblot Detection 10.10.1
Basic Protocol: Immunoprobing with Directly Conjugated Secondary Antibody 10.10.1
Alternate Protocol: Immunoprobing with Avidin-Biotin Coupling to Secondary Antibody 10.10.3
Support Protocol 1: Visualization with Chromogenic Substrates 10.10.5
Support Protocol 2: Visualization with Luminescent Substrates 10.10.6
Reagents and Solutions 10.10.7
Commentary 10.10.9

10.11 Autoradiography 10.11.1
Types of Film 10.11.1
Handling of Autoradiography Film 10.11.2
Quantitative Autoradiography 10.11.2
Exposure Procedures 10.11.2
Preflashing the Film 10.11.5
Film-Processing Recommendations 10.11.5
Literature Cited 10.11.6
Internet Resources 10.11.7

10.12 Digital Electrophoresis Analysis 10.12.1
Reasons for Digital Documentation and Analysis 10.12.1
Key Terms for Imaging 10.12.2
Image Capture 10.12.5
Analysis 10.12.8
Literature Cited 10.12.13

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CHAPTER 10
Electrophoresis

INTRODUCTION

Electrophoresis of protein samples in polyacrylamide gels is an indispensable analytical and, in some cases, preparative tool for the protein scientist. Electrophoresis can be used to separate and compare complex protein mixtures, evaluate purity of a protein during the course of its isolation, and provide estimates of physical characteristics such as subunit composition, isoelectric point, size, and charge.

Each type of electrophoretic separation can be conducted in a variety of gel sizes ranging from microgels (e.g., Phast gels from Hoefer Pharmacia) only slightly larger than a postage stamp to giant gels much larger than this page (Garrels, 1979; Young et al., 1983). In general, the time required to electrophorese, stain, and destain small gels is very short, and a minimal amount of sample is consumed. In contrast, larger gels consume more reagents, sample, and time, but provide increased resolution. Small gels are therefore recommended for rapid screening, and larger gels are indicated when maximum resolution is required, as in analysis of complex mixtures or samples containing very similar components.

The most common one-dimensional gel methods utilize the detergent sodium dodecyl sulfate (SDS) to solubilize, denature, and impart a strong negative charge to proteins (*UNIT 10.1*). Although many one-dimensional SDS gel methods have been published over the past 30 years, the single most widely used method is that initially described by Laemmli (1970). SDS-based gel separations are very robust as most, but not all, proteins are readily solubilized in SDS solutions. Most proteins bind a uniform amount of SDS per microgram of protein, which imparts a uniform charge density per unit mass to provide a separation based on the mass of the polypeptide chain.

An alternative one-dimensional separation methodology is based on the protein's isoelectric point and is usually performed in the presence of denaturants such as urea (*UNIT 10.2*). Like one-dimensional SDS gel separations, this denaturing isoelectric focusing method is capable of separating a large number of components in a single dimension. A third one-dimensional method uses native conditions to separate proteins based on intrinsic charge (*UNIT 10.3*). Although this method has less resolving power than the previous two methods, it can be particularly useful for associating a specific electrophoretically separated component with a biological activity, since no denaturants are used.

Two-dimensional electrophoresis (*UNIT 10.4*) involves the orthogonal combination of two different electrophoretic methods. Any two electrophoresis methods can be combined in this manner to produce useful separations. The most common two-dimensional method combines two individually high-resolution methods, consisting of isoelectric focusing (*UNIT 10.2*) followed by SDS gel electrophoresis (*UNIT 10.1*). Each of these separations individually is capable of resolving up to about 100 protein bands; when the two methods are combined, however, more than 1,000 proteins can be separated on a single two-dimensional gel (O'Farrell, 1975; Garrels, 1979). Hence this method can be used to analyze and compare complex protein mixtures, including whole-cell or tissue extracts. Other applications include assessment of the purity and homogeneity of a purified protein,

BOSTON CENTRAL UNIVERSITY LIBRARY

Contributed by David W. Speicher
Current Protocols in Protein Science (1998) 10.0.1-10.0.3
Copyright © 1998 by John Wiley & Sons, Inc.

including recombinant proteins (Chapter 7), and evaluation of some post-translational modifications (Chapters 12 and 13).

After electrophoretic separation, proteins are usually detected using general protein staining methods, such as Coomassie blue staining or the more sensitive silver staining. These staining methods "fix" the proteins in the gel matrix by either chemical cross-linking or denaturation to prevent subsequent diffusion of the protein bands (*UNIT 10.5*). When proteins are to be eluted or extracted from the gels for subsequent analyses, however, fixation-based stains should usually be avoided. In these cases, methods for protein detection without fixation should be used (*UNIT 10.6*) when it is necessary to stain a gel to locate the protein of interest.

An increasingly useful method of isolating proteins from gels for subsequent analysis involves electroblotting onto an inert support, such as polyvinylidene difluoride (PVDF) membranes (*UNIT 10.7*). As a result of the versatility of PVDF membranes, electroblotting has largely replaced electroelution as the method of choice for isolating proteins from most types of polyacrylamide gels (LeGendre and Matsudaira, 1988). PVDF membranes have very high protein binding capacities, good handling characteristics when either dry or wet, and are highly chemically inert. Hence they are compatible with many methods for subsequent analysis, including amino acid analysis, in situ chemical modification, in situ protease digestion, and N-terminal sequence analysis (Chapter 11). A wide range of general protein stains exist that are compatible with most blotting membranes (*UNIT 10.8*). Alternatively, radiolabeled proteins can be detected by autoradiography either in the gel or after transfer to a PVDF membrane (*UNIT 10.11*). The latter is normally preferable; as long as the proteins of interest are electrotransferred in high yield, autoradiography signals of weak β emitters such as ^{14}C or ^{35}S are usually higher on electroblotted membranes than in dried gels (even when fluorophores are impregnated in the gel), because the protein is primarily located on the surface of the membrane that was in contact with the gel during electroblotting. Historically, autoradiography was performed using medical X-ray films. More recently, a relatively wide range of commercially available autoradiography films and exposure methods have been developed that provide superior contrast and higher sensitivity for specific isotopes (*UNIT 10.11*).

The dimensional stability and chemical resistance of PVDF membranes provides an optimal medium for combining multiple detection methods, either sequentially on the same membrane, or on parallel lanes that can be reassembled after different detection methods have been performed. For example, sets of lanes containing replicate samples can be run on a single gel separated by prestained standards, which are used as guides to cut the resulting membrane. Each lane set can be subjected to different detection methods such as protein staining and immunoblotting with different antibodies (*UNIT 10.10*). The analyzed membranes can then be precisely reassembled to determine the relationships between bands detected by different methods. As there is no shrinkage or swelling of the membrane, this approach is particularly valuable for detecting slight differences in migration between components detected by different methods.

Capillary electrophoresis (CE) has recently emerged as a useful analytical method for separating peptides and, to a more limited extent, proteins (*UNIT 10.9*). In addition, CE protocols have been developed for preparative peptide isolation in adequate amounts for subsequent high sensitivity measurements such as micro-sequencing (Chapter 11) and mass spectrometry (Chapter 16). CE employs a fused-silica capillary column that may contain either free solution or a fluid matrix. In contrast to the relatively simple equipment requirements of polyacrylamide-based electrophoretic methods, commercial CE instruments are similar in complexity and expense to high-performance liquid chromatography (HPLC) instruments, due to the necessity for a high-voltage supply, a high-performance

BCU IAS CENTRAL UNIVERSITY LIBRARY

cooling system, and a very sensitive detector. Because CE separates components by charge rather than by hydrophobicity as in reversed-phase HPLC, it is a highly useful orthogonal separation method that is frequently used to complement HPLC peptide separation strategies. Compared with polyacrylamide methods, CE offers much higher resolution, flexibility, sensitivity, and speed for separating peptides (<6 kDa), whereas polyacrylamide methods are more generally applicable for protein separations.

Documentation of gel electrophoresis results for either archival or publication purposes historically used primarily silver halide–based photography. In recent years, the availability of economical hardware and software for capturing, storing, and analyzing digital images has led to the replacement of photography by digital analysis as the preferred documentation method (*UNIT 10.12*). Advantages of digital-based methods include more convenient storage and retrieval, inexpensive and rapid production of backup or additional copies, facile manipulation such as resizing and inserting into reports, and rapid data distribution to colleagues over the Internet. Digital processing of electrophoretic results is particularly essential for comparing large numbers of gels, developing databases of gel patterns and related information, and performing quantitative comparisons such as proteome studies utilizing two-dimensional gels.

LITERATURE CITED

Garrels, J.I. 1979. Two-dimensional gel electrophoresis and computer analysis of proteins synthesized by cloned cell lines. *J. Biol. Chem.* 54:7961-7977.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.

LeGendre, N. and Matsudaira, P. 1988. Direct protein microsequencing from Immobilon-P transfer membrane. *BioTechniques* 6:154-159.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.

Young, D.A., Voris, B.P., Maytin, E.V., and Colbert, R.A. 1983. Very high resolution two-dimensional electrophoretic separation of proteins on giant gels. *Methods Enzymol.* 91:190-214.

David W. Speicher

BCU IASI / CENTRAL UNIVERSITY LIBRARY

One-Dimensional SDS Gel Electrophoresis of Proteins

Electrophoresis is used to separate complex mixtures of proteins, to investigate subunit compositions, and to verify homogeneity of protein samples. It can also serve to purify proteins for use in further applications. In polyacrylamide gel electrophoresis, proteins migrate in response to an electrical field through pores in the gel matrix; pore size decreases with higher acrylamide concentrations. The combination of gel pore size and protein charge, size, and shape determines the migration rate of the protein.

Basic Protocol 1 gives the standard Laemmli method used for discontinuous gel electrophoresis under denaturing conditions, that is, in the presence of sodium dodecyl sulfate (SDS). In Basic Protocol 2, the standard method for full-size gels (e.g., 14 × 14 cm) is adapted for the minigel format (e.g., 7.3 × 8.3 cm). Minigels provide rapid separation but give lower resolution.

Several alternate protocols are provided for specific applications. The first and second alternate protocols cover electrophoresis of peptides and small proteins, separations that require modification of standard buffers; Alternate Protocol 1 uses a Tris-tricine buffer system, and Alternate Protocol 2 employs modified Tris buffers in the absence of urea. Alternate Protocol 3 describes continuous SDS-PAGE, a simplified method in which the same buffer is used for both gel and electrode solutions and the stacking gel is omitted. Other protocols cover the preparation and electrophoresis of various types of gels: ultrathin gels (Alternate Protocol 4), multiple single-concentration gels (Support Protocol 1), gradient gels (Alternate Protocol 5), multiple gradient gels (Support Protocol 2), and multiple gradient minigels (Support Protocol 3). Two-dimensional electrophoresis is discussed in *UNIT 10.4*.

CAUTION: Before any protocols are used, it is extremely important to read the following section about electricity and electrophoresis.

ELECTRICITY AND ELECTROPHORESIS

Many researchers are poorly informed concerning the electrical parameters of running a gel. It is important to note that the voltages and currents used during electrophoresis are dangerous and potentially lethal. Thus, safety should be an overriding concern. A working knowledge of electricity is an asset in determining what conditions to use and in troubleshooting the electrophoretic separation, if necessary. For example, an unusually high or low voltage for a given current (milliampere) might indicate an improperly made buffer or an electrical leak in the chamber.

Safety Considerations

1. Never remove or insert high-voltage leads unless the power supply voltage is turned down to zero and the power supply is turned off. Always grasp high-voltage leads one at a time with one hand only. Never insert or remove high-voltage leads with both hands. This can shunt potentially lethal electricity through the chest and heart should electrical contact be made between a hand and a bare wire. On older or homemade instruments, the banana plugs may not be shielded and can still be connected to the power supply at the same time they make contact with a hand. Carefully inspect all cables and connections and replace frayed or exposed wires immediately.

2. Always start with the power supply turned off. Have the power supply controls turned all the way down to zero. Then hook up the gel apparatus: generally, connect the red

BCM&S CENTRAL UNIVERSITY LIBRARY

Copyright © 1995 by John Wiley & Sons, Inc.

high-voltage lead to the red outlet and the black high-voltage lead to the black outlet. Turn the power supply on with the controls set at zero and the high-voltage leads connected. Then, turn up the voltage, current, or power to the desired level. Reverse the process when the power supply is turned off: i.e., to disconnect the gel, turn the power supply down to zero, wait for the meters to read zero, turn off the power supply, and then disconnect the gel apparatus one lead at a time.

CAUTION: *If the gel is first disconnected and then the power supply turned off, a considerable amount of electrical charge is stored internally. The charge will stay in the power supply over a long time. This will discharge through the outlets even though the power supply is turned off and can deliver an electrical shock.*

Ohm's Law and Electrophoresis

Understanding how a gel apparatus is connected to the power supply requires a basic understanding of Ohm's law: voltage = current × resistance, or $V = IR$. A gel can be viewed as a resistor and the power supply as the voltage and current source. Most power supplies deliver constant current or constant voltage. Some will also deliver constant power: power = voltage × current, or $VI = I^2R$. The discussion below focuses on constant current because this is the most common mode in vertical SDS-PAGE.

Most modern commercial equipment is color-coded so that the red or positive terminal of the power supply can simply be connected to the red lead of the gel apparatus, which goes to the lower buffer chamber. The black lead is connected to the black or negative terminal and goes to the upper buffer chamber. This configuration is designed to work with vertical slab gel electrophoreses in which negatively charged proteins or nucleic acids move to the positive electrode in the lower buffer chamber (an anionic system).

When a single gel is attached to a power supply, the negative charges flow from the negative cathode (black) terminal into the upper buffer chamber, through the gel, and into the lower buffer chamber. The lower buffer chamber is connected to the positive anode (red) terminal to complete the circuit. Thus, negatively charged molecules, such as SDS-coated proteins and nucleic acids, move from the negative cathode attached to the upper buffer chamber toward the positive anode attached to the lower chamber. SDS-PAGE is an anionic system because of the negatively charged SDS.

Occasionally, proteins are separated in cationic systems. In these gels, the proteins are positively charged because of the very low pH of the gel buffers (e.g., acetic acid/urea gels for histone separations) or the presence of a cationic detergent (e.g., cetyltrimethylammonium bromide, CTAB). Proteins move toward the negative electrode (cathode) in cationic gel systems, and the polarity is reversed compared to SDS-PAGE: the red lead from the lower buffer chamber is attached to the black outlet of the power supply, and the black lead from the upper buffer chamber is attached to the red outlet of the power supply.

Most SDS-PAGE separations are performed under constant current (consult manufacturer's instructions to set the power supply for constant current operation). The resistance of the gel will increase during SDS-PAGE in the standard Laemmli system. If the current is constant, then the voltage will increase during the run as the resistance goes up.

Power supplies usually have more than one pair of outlets. The pairs are connected in parallel with one another internally. If more than one gel is connected directly to the outlets of a power supply, then these gels are connected in parallel. In a parallel circuit, the voltage is the same across each gel. In other words, if the power supply reads 100 V, then each gel has 100 V across its electrodes. The total current, however, is the sum of the individual currents going through each gel. Therefore, under constant current it is necessary to

RCU IAS/CENTRAL UNIVERSITY LIBRARY

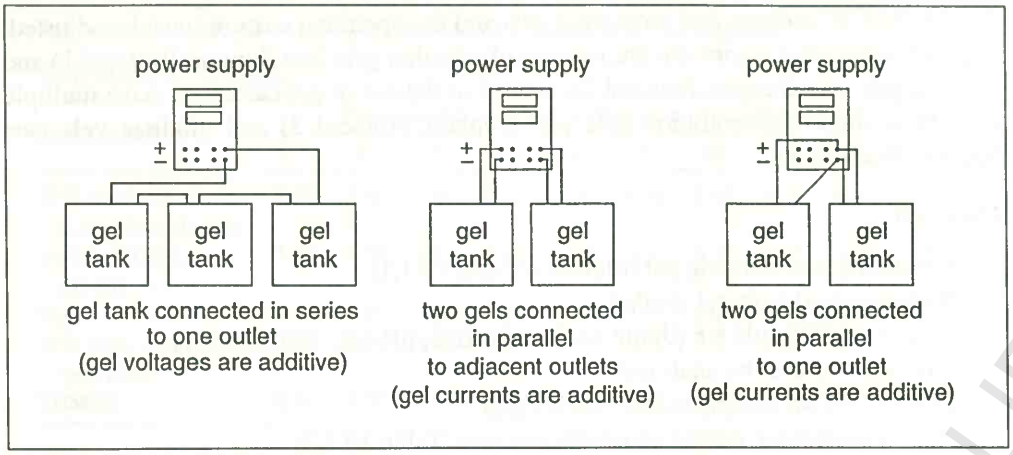

power supply
power supply
power supply
gel tank
gel tank
gel tank
gel tank
gel tank
gel tank
gel tank
gel tank
gel tank connected in series to one outlet (gel voltages are additive)
two gels connected in parallel to adjacent outlets (gel currents are additive)
two gels connected in parallel to one outlet (gel currents are additive)

Figure 10.1.1 Series and parallel connections of gel tanks to power supply.

increase the current for each additional gel that is connected to the power supply. Two identical gels require double the current to achieve the same starting voltages and electrophoresis separation times.

Multiple gel apparatuses can also be connected to one pair of outlets on a power supply. This is useful with older power supplies that have a limited number of outlets. When connecting several gel units to one outlet, make certain the connections between the units are shielded and protected from moisture. The gels can be connected in parallel or in series (Fig. 10.1.1). In the case of two or more gels running off the same outlet in series, the current is the same for every gel. If 10 mA is displayed by the power supply meter, for example, each gel in series will experience 10 mA. The voltage, however, is additive for each gel. If one gel at a constant 10 mA produces 100 V, then two identical gels in series will produce 200 V (100 V each) and so on. Thus, the voltage can limit the number of units connected in series on low-voltage power supplies.

Gel thickness affects the above relationships. A 1.5-mm gel can be thought of as consisting of two 0.75-mm-thick gels run in parallel. Because currents are additive in parallel circuits, a 0.75-mm gel will require half the current of the 1.5-mm gel to achieve the same starting voltage and separation time. If a gel thickness is doubled, then the current must also be doubled. There are limits to the amount of current that can be applied. Thicker gels require more current, generating more heat that must be dissipated. Unless temperature control is available in the gel unit, a thick gel should be run more slowly than a thin gel.

NOTE: Milli-Q-purified water or equivalent should be used throughout the protocols.

DENATURING (SDS) DISCONTINUOUS GEL ELECTROPHORESIS: LAEMMLI GEL METHOD

BASIC PROTOCOL 1

One-dimensional gel electrophoresis under denaturing conditions (i.e., in the presence of 0.1% SDS) separates proteins based on molecular size as they move through a polyacrylamide gel matrix toward the anode. The polyacrylamide gel is cast as a separating gel (sometimes called resolving or running gel) topped by a stacking gel and secured in an electrophoresis apparatus. After sample proteins are solubilized by boiling in the presence of SDS, an aliquot of the protein solution is applied to a gel lane, and the individual proteins are separated electrophoretically. 2-Mercaptoethanol (2-ME) or dithiothreitol (DTT) is added during solubilization to reduce disulfide bonds.

This protocol is designed for a vertical slab gel with a maximum size of 0.75 mm × 14 cm × 14 cm. For thicker gels, or minigels (see Basic Protocol 2 and Support Protocol 5),

BOSTON COLLEGE UNIVERSITY LIBRARY

the volumes of stacking and separating gels and the operating current must be adjusted. Support protocols describe the preparation of ultrathin gels (see Support Protocol 1) and gradient gels (see Support Protocol 3), as well as the use of gel casters to make multiple gels, both single-concentration gels (see Support Protocol 2) and gradient gels (see Support Protocol 4).

Materials

Separating and stacking gel solutions (Table 10.1.1)
H_2O-saturated isobutyl alcohol
1× Tris·Cl/SDS, pH 8.8 (dilute 4× Tris·Cl/SDS, pH 8.8; Table 10.1.1)
Protein sample to be analyzed
2× and 1× SDS sample buffer (see recipe)
Protein molecular-weight-standards mixture (Table 10.1.2)
6× SDS sample buffer (see recipe; optional)
1× SDS electrophoresis buffer (see recipe)

Electrophoresis apparatus: Protean II 16-cm cell (Bio-Rad) or SE 600/400 16-cm unit (Hoefer Pharmacia) with clamps, glass plates, casting stand, and buffer chambers
0.75-mm spacers
0.45-µm filters (used in stock solution preparation)
25-ml Erlenmeyer side-arm flask
Vacuum pump with cold trap
0.75-mm Teflon comb with 1, 3, 5, 10, 15, or 20 teeth
25- or 100-µl syringe with flat-tipped needle
Constant-current power supply (see introduction)

Pour the separating gel

1. Assemble the glass-plate sandwich of the electrophoresis apparatus according to manufacturer's instructions using two clean glass plates and two 0.75-mm spacers.

   If needed, clean the glass plates in liquid Alconox or RBS-35 (Pierce). These aqueous-based solutions are compatible with silver and Coomassie blue staining procedures.

2. Lock the sandwich to the casting stand.

3. Prepare the separating gel solution as directed in Table 10.1.1, degassing using a rubber-stoppered 25-ml Erlenmeyer side-arm flask connected with vacuum tubing to a vacuum pump with a cold trap. After adding the specified amount of 10% ammonium persulfate and TEMED to the degassed solution, stir gently to mix.

   Table 10.1.1 was prepared as a convenient summary to aid in the preparation of separating and stacking gels. The stacking gel is the same regardless of the separating gel used.

   The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. Generally, use 5% gels for SDS-denatured proteins of 60 to 200 kDa, 10% gels for SDS-denatured proteins of 16 to 70 kDa, and 15% gels for SDS-denatured proteins of 12 to 45 kDa (Table 10.1.1).

4. Using a Pasteur pipet, apply separating gel solution to the sandwich along an edge of one of the spacers until the height of the solution between the glass plates is ~11 cm.

   Use the solution immediately; otherwise it will polymerize in the flask.

   Sample volumes <10 µl do not require a stacking gel. In this case, cast the resolving gel as you normally would, but extend the resolving gel into the comb to form the well. The proteins are then separated under the same conditions as used when a stacking gel is present. Although this protocol works well with single-concentration gels, a gradient gel is recommended for maximum resolution (see Alternate Protocol 5).

PGU IASLIC-PEYRA UNIVERSITY LIBRARY

Table 10.1.1 Recipes for Polyacrylamide Separating and Stacking Gels[a]

SEPARATING GEL

Stock solution[b]	Final acrylamide concentration in separating gel (%)[c]									
	5	6	7	7.5	8	9	10	12	13	15
30% acrylamide/ 0.8% bisacrylamide	2.50	3.00	3.50	3.75	4.00	4.50	5.00	6.00	6.50	7.50
4× Tris·Cl/SDS, pH 8.8	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H_2O	8.75	8.25	7.75	7.50	7.25	6.75	6.25	5.25	4.75	3.75
10% (w/v) ammonium persulfate[d]	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
TEMED	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Preparation of separating gel

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution, 4× Tris·Cl/SDS, pH 8.8 (see reagents, below), and H_2O. Degas under vacuum about 5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

STACKING GEL (*3.9% acrylamide*)

In a 25-ml side-arm flask, mix 0.65 ml of 30% acrylamide/0.8% bisacrylamide, 1.25 ml of 4× Tris·Cl/SDS, pH 6.8 (see reagents, below), and 3.05 ml H_2O. Degas under vacuum 10 to 15 min. Add 25 µl of 10% ammonium persulfate and 5 µl TEMED. Swirl gently to mix. Use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

REAGENTS USED IN GELS

30% acrylamide/0.8% bisacrylamide

Mix 30.0 g acrylamide and 0.8 g *N,N'*-methylenebisacrylamide with H_2O in a total volume of 100 ml. Filter the solution through a 0.45-µm filter and store at 4°C in the dark. The 2× crystallized grades of acrylamide and bisacrylamide are recommended. Discard after 30 days, as acrylamide gradually hydrolyzes to acrylic acid and ammonia.

CAUTION: *Acrylamide monomer is neurotoxic. Mask should be worn when weighing acrylamide powder. Gloves should be worn while handling the solution, and the solution should not be pipetted by mouth.*

4× Tris·Cl/SDS, pH 6.8 (0.5 M Tris·Cl containing 0.4% SDS)

Dissolve 6.05 g Tris base in 40 ml H_2O. Adjust to pH 6.8 with 1 N HCl. Add H_2O to 100 ml total volume. Filter the solution through a 0.45-µm filter, add 0.4 g SDS, and store at 4°C up to 1 month.

4× Tris·Cl/SDS, pH 8.8 (1.5 M Tris·Cl containing 0.4% SDS)

Dissolve 91 g Tris base in 300 ml H_2O. Adjust to pH 8.8 with 1 N HCl. Add H_2O to 500 ml total volume. Filter the solution through a 0.45-µm filter, add 2 g SDS, and store at 4°C up to 1 month.

[a]The recipes produce 15 ml of separating gel and 5 ml of stacking gel, which are adequate for a gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the SDS (denaturing) discontinuous buffer system of Laemmli (1970).
[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.
[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. See annotation to step 3, Basic Protocol 1.
[d]Best to prepare fresh.

RESTRICTED - MSU RETTER UNIVERSITY LIBRARY

5. Using another Pasteur pipet, slowly cover the top of the gel with a layer (~1 cm thick) of H_2O-saturated isobutyl alcohol, by gently layering the isobutyl alcohol against the edge of one and then the other of the spacers.

Be careful not to disturb the gel surface. The overlay provides a barrier to oxygen, which inhibits polymerization, and allows a flat interface to form during gel formation.

The H_2O-saturated isobutyl alcohol is prepared by shaking isobutyl alcohol and H_2O in a separatory funnel. The aqueous (lower) phase is removed. This procedure is repeated several times. The final upper phase is H_2O-saturated isobutyl alcohol.

6. Allow the gel to polymerize 30 to 60 min at room temperature.

A sharp optical discontinuity at the overlay/gel interface will be visible on polymerization. Failure to form a firm gel usually indicates a problem with the ammonium persulfate, TEMED (N,N,N′,N′-tetramethylethylenediamine), or both. Ammonium persulfate solution should be made fresh before use. Ammonium persulfate should "crackle" when added to the water. If not, fresh ammonium persulfate should be purchased. Purchase TEMED in small bottles so, if necessary, a new previously unopened source can be tried.

Pour the stacking gel

7. Pour off the layer of H_2O-saturated isobutyl alcohol and rinse with 1× Tris·Cl/SDS, pH 8.8.

Residual isobutyl alcohol can reduce resolution of the protein bands; therefore, it must be completely removed. The isobutyl alcohol overlay should not be left on the gel longer than 2 hr.

8. Prepare the stacking gel solution as directed in Table 10.1.1.

Use the solution immediately to keep it from polymerizing in the flask.

9. Using a Pasteur pipet, slowly allow the stacking gel solution to trickle into the center of the sandwich along an edge of one of the spacers until the height of the solution in the sandwich is ~1 cm from the top of the plates.

Be careful not to introduce air bubbles into the stacking gel.

10. Insert a 0.75-mm Teflon comb into the layer of stacking gel solution. If necessary, add additional stacking gel to fill the spaces in the comb completely.

Again, be careful not to trap air bubbles in the tooth edges of the comb; they will cause small circular depressions in the well after polymerization that will lead to distortion in the protein bands during separation.

11. Allow the stacking gel solution to polymerize 30 to 45 min at room temperature.

A sharp optical discontinuity will be visible around wells on polymerization.

Prepare the sample and load the gel

12. Dilute a portion of the protein sample to be analyzed 1:1 (v/v) with 2× SDS sample buffer and heat 3 to 5 min at 100°C in a sealed screw-cap microcentrifuge tube. If the sample is a precipitated protein pellet, dissolve the protein in 50 to 100 µl of 1× SDS sample buffer and boil 3 to 5 min at 100°C. Dissolve protein-molecular-weight standards mixture in 1× SDS sample buffer according to supplier's instructions as a control (Table 10.1.2).

For dilute protein solutions, consider adding 5:1 protein solution/6× SDS sample buffer to increase the amount of protein loaded. Proteins can also be concentrated by precipitation in acetone, ethanol, or trichloroacetic acid (TCA), but losses will occur.

PCU IAS / CENTRAL UNIVERSITY LIBRARY

Table 10.1.2 Molecular Weights of Protein Standards for Polyacrylamide Gel Electrophoresis[a]

Protein	Molecular weight
Cytochrome *c*	11,700
α-Lactalbumin	14,200
Lysozyme (hen egg white)	14,300
Myoglobin (sperm whale)	16,800
β-Lactoglobulin	18,400
Trypsin inhibitor (soybean)	20,100
Trypsinogen, PMSF treated	24,000
Carbonic anhydrase (bovine erythrocytes)	29,000
Glyceraldehyde-3-phosphate dehydrogenase (rabbit muscle)	36,000
Lactate dehydrogenase (porcine heart)	36,000
Aldolase	40,000
Ovalbumin	45,000
Catalase	57,000
Bovine serum albumin	66,000
Phosphorylase *b* (rabbit muscle)	97,400
β-Galactosidase	116,000
RNA polymerase, *E. coli*	160,000
Myosin, heavy chain (rabbit muscle)	205,000

[a]Protein standards are commercially available in kits (e.g., Hoefer Pharmacia Biotech, GIBCO/BRL, Bio-Rad, or Sigma).

For a 0.8-cm-wide well, 25 to 50 μg total protein in <20 μl is recommended for a complex mixture when staining with Coomassie blue, and 1 to 10 μg total protein is needed for samples containing one or a few proteins. If silver staining is used, 10- to 100-fold less protein can be applied (0.01 to 5 μg in <20 μl depending on sample complexity).

To achieve the highest resolution possible, the following precautions are recommended. Prior to adding the sample buffer, keep samples at 0°C. Add the SDS sample buffer (room temperature) directly to the 0°C sample (still on ice) in a screw-top microcentrifuge tube. Cap the tube to prevent evaporation, vortex, and transfer directly to a 100°C water bath for 3 to 5 min. Let immunoprecipitates dissolve for 1 hr at 56°C in 1× SDS sample buffer prior to boiling. DO NOT leave the sample in SDS sample buffer at room temperature without first heating to a 100°C to inactivate proteases (see Critical Parameters and Troubleshooting). Endogenous proteases are very active in SDS sample buffer and will cause severe degradation of the sample proteins after even a few minutes at room temperature. To test for possible proteases, mix the sample with SDS sample buffer without heating and leave at room temperature for 1 to 3 hr. A loss of high-molecular-weight bands and a general smearing of the banding pattern indicate a protease problem. Once heated, the samples can sit at room temperature for the time it takes to load samples.

13. Carefully remove the Teflon comb without tearing the edges of the polyacrylamide wells. After the comb is removed, rinse wells with 1× SDS electrophoresis buffer.

    The rinse removes unpolymerized monomer; otherwise, the monomer will continue to polymerize after the comb is removed, creating uneven wells that will interfere with sample loading and subsequent separation.

14. Using a Pasteur pipet, fill the wells with 1× SDS electrophoresis buffer.

BOURSE CENTRAL UNIVERSITY LIBRARY

If well walls are not upright, they can be manipulated with a flat-tipped needle attached to a syringe.

15. Attach gel sandwich to upper buffer chamber using manufacturer's instructions.
16. Fill lower buffer chamber with the recommended amount of 1× SDS electrophoresis buffer.
17. Place sandwich attached to upper buffer chamber into lower buffer chamber.
18. Partially fill the upper buffer chamber with 1× SDS electrophoresis buffer so that the sample wells of the stacking gel are filled with buffer.

    Monitor the upper buffer chamber for leaks and if necessary, reassemble the unit. A slow leak in the upper buffer chamber may cause arcing around the upper electrode and damage the upper buffer chamber.

19. Using a 25- or 100-µl syringe with a flat-tipped needle, load the protein sample(s) into one or more wells by carefully applying the sample as a thin layer at the bottom of the wells. Load control wells with molecular weight standards. Add an equal volume of 1× SDS sample buffer to any empty wells to prevent spreading of adjoining lanes.

    Preparing the samples at approximately the same concentration and loading an equal volume to each well will ensure that all lanes are the same width and that the proteins run evenly. If unequal volumes of sample buffer are added to wells, the lane with the larger volume will spread during electrophoresis and constrict the adjacent lanes, causing distortions.

    The samples will layer on the bottom of the wells because the glycerol added to the sample buffer gives the solution a greater density than the electrophoresis buffer. The bromphenol blue in the sample buffer makes sample application easy to follow visually.

20. Fill the remainder of the upper buffer chamber with additional 1× SDS electrophoresis buffer so that the upper platinum electrode is completely covered. Do this slowly so that samples are not swept into adjacent wells.

Run the gel

21. Connect the power supply to the cell and run at 10 mA of constant current for a slab gel 0.75 mm thick, until the bromphenol blue tracking dye enters the separating gel. Then increase the current to 15 mA.

    For a standard 16-cm gel sandwich, 4 mA per 0.75-mm-thick gel will run ~15 hr (i.e., overnight); 15 mA per 0.75-mm gel will take 4 to 5 hr. To run two gels or a 1.5-mm-thick gel, simply double the current. When running a 1.5-mm gel at 30 mA, the temperature must be controlled (10° to 20°C) with a circulating constant-temperature water bath to prevent "smiling" (curvature in the migratory band). Temperatures <5°C should not be used because SDS in the running buffer will precipitate. If the level of buffer in the upper chamber decreases, a leak has occurred.

22. After the bromphenol blue tracking dye has reached the bottom of the separating gel, disconnect the power supply.

    See Safety Considerations in the introduction.

Disassemble the gel

23. Discard electrode buffer and remove the upper buffer chamber with the attached gel sandwich.
24. Orient the gel so that the order of the sample wells is known, remove the sandwich from the upper buffer chamber, and lay the sandwich on a sheet of absorbent paper or paper towels.

RCU IASL / KENTRAL UNIVERSITY LIBRARY

25. Carefully slide one of the spacers halfway from the edge of the sandwich along its entire length. Use the exposed spacer as a lever to pry open the glass plate, exposing the gel.

26. Carefully remove the gel from the lower plate. Cut a small triangle off one corner of the gel so the lane orientation is not lost during staining and drying. Proceed with protein detection.

The gel can be stained with Coomassie blue or silver (UNIT 10.5), or proteins can be electroeluted, electroblotted onto a polyvinylidene difluoride (PVDF) membrane for subsequent staining or sequence analysis (UNIT 10.7), or transferred to a membrane for immunoblotting. If the proteins are radiolabeled (UNIT 3.3), they can be detected by autoradiography.

ELECTROPHORESIS IN TRIS-TRICINE BUFFER SYSTEMS

ALTERNATE PROTOCOL 1

Separation of peptides and proteins under 10 to 15 kDa is not possible in the traditional Laemmli discontinuous gel system (Basic Protocol 1). This is due to the comigration of SDS and smaller proteins, obscuring the resolution. Two approaches to obtain the separation of small proteins and peptides in the range of 5 to 20 kDa are presented: this Tris-tricine method that follows and a system using increased buffer concentrations (see Alternate Protocol 2). The Tris-tricine system uses a modified buffer to separate the SDS and peptides, thus improving resolution. Several precast gels are available for use with the tricine formulations (Table 10.1.3).

Additional Materials (also see Basic Protocol 1)

Separating and stacking gel solutions (Table 10.1.4)
2× tricine sample buffer (see recipe)
Peptide molecular-weight-standards mixture (Table 10.1.5)
Cathode buffer (see recipe)
Anode buffer (see recipe)
Coomassie blue G-250 staining solution (see recipe)
10% (v/v) acetic acid

Table 10.1.3 Vertical Format Precast Gel Compatibility

Gel type and compatibility	Gel supplier				
	Bio-Rad	ISS/Daiichi	Jule	Millipore	Novex
SDS-PAGE gel type offered					
Peptide (tricine)	×	×	×	×	×
Single concentration	×	×	×	×	×
Gradient	×	×	×	×	×
Minigel size	×	×	×	×	×
Standard gel size		×	×		
Compatibility of gel with equipment manufactured by					
Hoefer Pharmacia Biotech		×	×	×	×
Bio-Rad	×	×	×	×	
GIBCO/BRL	×	×	×	×	×
Novex		×		×	×
ISS/Daiichi		×	×	×	

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

Table 10.1.4 Recipes for Tricine Peptide Separation Gels[a]

SEPARATING AND STACKING GELS

Stock solution[b]	Separating gel	Stacking gel
30% acrylamide/0.8% bisacrylamide	9.80 ml	1.62 ml
Tris·Cl/SDS, pH 8.45	10.00 ml	3.10 ml
H_2O	7.03 ml	7.78 ml
Glycerol	4.00 g (3.17 ml)	—
10% (w/v) ammonium persulfate[c]	50 µl	25 µl
TEMED	10 µl	5 µl

Prepare separating and stacking gel solutions separately.

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (Table 10.1.1), Tris·Cl/SDS, pH 8.45 (see reagents, below), and H_2O. Add glycerol to separating gel only. Degas under vacuum 10 to 15 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix, use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

ADDITIONAL REAGENTS USED IN GELS

Tris·Cl/SDS, pH 8.45 (3.0 M Tris·Cl 0.3% SDS)

Dissolve 182 g Tris base in 300 ml H_2O. Adjust to pH 8.45 with 1 N HCl. Add H_2O to 500 ml total volume. Filter the solution through a 0.45-µm filter, add 1.5 g SDS, and store at 4°C up to 1 month.

[a]The recipes produce 30 ml of separating gel and 12.5 ml of stacking gel, which are adequate for two gels of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the Tris-tricine buffer system of Schagger and von Jagow (1987).

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Best to prepare fresh.

1. Prepare and pour the separating and stacking gels (see Basic Protocol 1, steps 1 to 11), using Table 10.1.4 in place of Table 10.1.1.
2. Prepare the sample (see Basic Protocol 1, step 12), but make the following changes for tricine gels. Substitute 2× tricine sample buffer for the 2× SDS sample buffer. Dilute an aliquot of the protein or peptide sample to be analyzed 1:1 (v/v) with 2× tricine sample buffer. Treat the sample at 40°C for 30 to 60 min prior to loading.

   If proteolytic activity is a problem, heating samples to 100°C for 3 to 5 min before loading the wells may be required (see Basic Protocol 1, annotation to step 12). Use the peptide molecular-weight-standards mixture for peptide separations (Table 10.1.5).
3. Load the gel and set up the electrophoresis apparatus (see Basic Protocol 1, steps 13 to 20) with the following alterations. Remove comb and, using the tricine-containing cathode buffer or water, rinse once and fill wells. Fill the lower buffer chamber with anode buffer, assemble the unit, and attach the upper buffer chamber. Fill the upper buffer chamber with cathode buffer and load the samples.
4. Connect the power supply to the cell and run 1 hr at 30 V (constant voltage) followed by 4 to 5 hr at 150 V (constant voltage). Use heat exchanger to keep the electrophoresis chamber at room temperature.

PGU IASL CENTRAL UNIVERSITY LIBRARY

Table 10.1.5 Molecular Weights of Peptide Standards for Polyacrylamide Gel Electrophoresis[a]

Peptide	Molecular weight
Myoglobin (polypeptide backbone)	16,950
Myoglobin 1-131	14,440
Myoglobin 56-153	10,600
Myoglobin 56-131	8,160
Myoglobin 1-55	6,210
Glucagon	3,480
Myoglobin 132-153	2,510

[a]Peptide standards are commercially available from Sigma.

5. After the tracking dye has reached the bottom of the separating gel, disconnect the power supply.

   See Safety Considerations in the introduction.

   Coomassie blue G-250 is used as a tracking dye instead of bromphenol blue because it moves ahead of the smallest peptides.

6. Disassemble the gel (see Basic Protocol 1, steps 23 to 26). Stain proteins in the gel for 1 to 2 hr in Coomassie blue G-250 staining solution. Follow by destaining with 10% acetic acid, changing the solution every 30 min until background is clear (3 to 5 changes). For higher sensitivity, use silver staining as a recommended alternative (*UNIT 10.5*).

   Prolonged staining and destaining will result in the loss of resolution of the smaller proteins (<10 kDa). Proteins diffuse within the gel and out of the gel, resulting in a loss of staining intensity and resolution.

NONUREA PEPTIDE SEPARATIONS WITH TRIS BUFFERS

ALTERNATE PROTOCOL 2

A simple modification of the traditional Laemmli buffer system presented in Basic Protocol 1, in which the increased concentration of buffers provides better separation between the stacked peptides and the SDS micelles, permits reasonable separation of peptides as small as 5 kDa.

Additional Materials (also see Basic Protocol 1)

Separating and stacking gel solutions (Table 10.1.6)
2× SDS electrophoresis buffer (see recipe)
2× Tris·Cl/SDS, pH 8.8 (dilute 4× Tris·Cl/SDS, pH 8.8; Table 10.1.1)

1. Prepare and pour the separating gel (see Basic Protocol 1, steps 1 to 6) using Table 10.1.6 in place of Table 10.1.1.

2. Prepare and pour the stacking gel (see Basic Protocol 1, steps 7 to 11), using 2× Tris·Cl/SDS, pH 8.8, rather than 1× Tris·Cl/SDS buffer, for rinsing the gel after removing the isobutyl alcohol overlay.

3. Prepare the sample and load the gel (see Basic Protocol 1, steps 12 to 20) and substitute 2× SDS electrophoresis buffer for the 1× SDS electrophoresis buffer.

   Table 10.1.5 lists the standards for small protein separations.

BOURNS / CENTRAL UNIVERSITY LIBRARY

Table 10.1.6 Recipes for Modified Laemmli Peptide Separation Gels[a]

SEPARATING AND STACKING GELS

Stock solution[b]	Separating gel	Stacking gel
30% acrylamide/0.8% bisacrylamide	10.00 ml	0.65 ml
8× Tris·Cl, pH 8.8	3.75 ml	—
4× Tris·Cl, pH 6.8	—	1.25 ml
10% (w/v) SDS	0.15 ml	50 µl
H_2O	1.00 ml	3.00 ml
10% (w/v) ammonium persulfate[c]	50 µl	25 µl
TEMED	10 µl	5 µl

Prepare separating and stacking gel solutions separately.

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 8× Tris·Cl, pH 8.8 (separating gel) or 4× Tris·Cl, pH 6.8 (stacking gel), 10% SDS (see reagents, below), and H_2O. Degas under vacuum 10 to 15 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix; use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

ADDITIONAL REAGENTS USED IN GELS

4× Tris·Cl, pH 6.8 (0.5 M Tris·Cl)

Dissolve 6.05 g Tris base in 40 ml H_2O. Adjust to pH 6.8 with 1 N HCl. Add H_2O to 100 ml total volume. Filter the solution through a 0.45-µm filter and store at 4°C up to 1 month.

8× Tris·Cl, pH 8.8 (3.0 M Tris·Cl)

Dissolve 182 g Tris base in 300 ml H_2O. Adjust to pH 8.8 with 1 N HCl. Add H_2O to 500 ml total volume. Filter the solution through a 0.45-µm filter and store at 4°C up to 1 month.

10% (w/v) SDS

Mix 1 g SDS in 10 ml of H_2O. Use immediately.

[a]The recipes produce 15 ml of separating gel and 5 ml of stacking gel, which are adequate for one gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the modified Laemmli peptide separation system of Okajima et al. (1993).

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Best to prepare fresh.

4. Run the gel (see Basic Protocol 1, steps 21 and 22).

Note that the separations will take ~25% longer than those using Basic Protocol 1. The increased buffer concentrations lead to faster transit through the stacking gel but lower mobility in the resolving gel.

5. Disassemble the gel (see Basic Protocol 1, steps 23 to 26).

If Coomassie blue staining is used, stain using the solutions outlined in UNIT 10.5 for 1 hr followed by destaining with several changes of 25% methanol/7% acetic acid. Prolonged staining will result in the loss of the smaller proteins.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CONTINUOUS SDS-PAGE

With continuous SDS-PAGE, the same buffer is used for both the gel and electrode solutions. Although continuous gels lack the resolution of the discontinuous systems, they are extremely versatile, less prone to mobility artifacts, and much easier to prepare. The stacking gel is omitted.

Additional Materials (also see Basic Protocol 1)

Separating gel solution (Table 10.1.7)
2× and 1× phosphate/SDS sample buffer (see recipe)
1× phosphate/SDS electrophoresis buffer (see recipe)

1. Prepare and pour a single separating gel (see Basic Protocol 1, steps 1 to 4), except use solutions in Table 10.1.7 in place of those in Table 10.1.1 and fill the gel sandwich to the top. Omit the stacking gel. Insert the comb (see Basic Protocol 1, step 10) and allow the gel to polymerize 30 to 60 min at room temperature.

Table 10.1.7 Recipes for Separating Gels for Continuous SDS-PAGE[a]

SEPARATING GEL

Stock solution[b]	Final acrylamide concentration in separating gel (%)[c]										
	5	6	7	8	9	10	11	12	13	14	15
30% acrylamide/ 0.8% bisacrylamide	2.5	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50	7.00	7.50
4× phosphate/SDS, pH 7.2	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H_2O	8.75	8.25	7.75	7.25	6.75	6.25	5.75	5.25	4.75	4.25	3.75
10% (w/v) ammonium persulfate[d]	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
TEMED	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Preparation of separating gel

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 4× phosphate/SDS, pH 7.2, and H_2O. Degas under vacuum about 5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

ADDITIONAL REAGENTS USED IN GELS

4× phosphate/SDS, pH 7.2 (0.4 M sodium phosphate/0.4% SDS)

Mix 46.8 g $NaH_2PO_4 \cdot H_2O$, 231.6 g $Na_2HPO_4 \cdot 7\ H_2O$, and 12 g SDS in 3 liters H_2O.

Store at 4°C for up to 3 months.

[a]The recipes produce 15 ml of separating gel, which is adequate for one gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the original continuous phosphate buffer system of Weber et al. (1972). The stacking gel is omitted.

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. See Basic Protocol 1, annotation to step 3.

[d]Best to prepare fresh.

BOSTON CENTRAL UNIVERSITY LIBRARY

2. Mix the protein sample 1:1 with 2× phosphate/SDS sample buffer and heat to 100°C for 2 min.

   For large sample volumes or samples suspended in high ionic strength buffers (>50 mM), dialyze against 1× sample buffer prior to electrophoresis. Note that the precautions about proteases (see Basic Protocol 1, step 12) apply here.

3. Assemble the electrophoresis apparatus and load the sample (see Basic Protocol 1, steps 13 to 20), using the phosphate/SDS electrophoresis buffer and loading empty wells with 1× phosphate/SDS sample buffer.

4. Connect the power supply and start the run with 15 mA per 0.75-mm-thick gel until the tracking dye has entered the gel. Continue electrophoresis at 30 mA for 3 hr (5% gel), 5 hr (10% gel), 8 hr (15% gel), or until the dye reaches the bottom of the gel.

   Use temperature control if available to maintain the gel at 15° to 20°C. SDS will precipitate below 15°C in this system.

5. Disassemble the gel (see Basic Protocol 1, steps 23 to 26).

   Proteins in the gel may be stained using the protocols given in UNIT 10.5.

ALTERNATE PROTOCOL 4

CASTING AND RUNNING ULTRATHIN GELS

Ultrathin gels provide superb resolution but are difficult to handle. In this application, gels are cast on Gel Bond, a Mylar support material. Silver staining is recommended for the best resolution. Combs and spacers for gels <0.5 mm thick are not readily available for most protein electrophoresis units. However, by adapting combs and spacers used for DNA sequencing, casting gels from 0.2 to 0.5 mm thick is straightforward.

Additional Materials *(also see Basic Protocol 1)*

95% (v/v) ethanol
Gel Bond (FMC) cut to a size slightly smaller than the gel plate dimensions
Glue stick
Ink roller (available from art supply stores)
Combs and spacers (0.19 to 0.5 mm; sequencing gel spacers and combs can be cut to fit)

1. Wash gel plates with water-based laboratory detergent followed by successive rinses with hot tap water, deionized water, and finally 95% ethanol. Allow to air dry.

   Gel plates must be extremely clean for casting thin gels.

2. Apply a streak of adhesive from a glue stick to the bottom edge of the glass plate. Quickly position the Gel Bond with the hydrophobic side down (a drop of water will bead up on the hydrophobic surface). Apply pressure with Kimwipe tissue to attach the Gel Bond firmly to the plate. Last, pull the top portion of the Gel Bond back, place a few drops of water underneath, and roll flat with an ink roller.

   Make sure the Gel Bond does not extend beyond the edges of the upper and lower sealing surface of the plate. This will cause it to buckle on sealing. Reposition the Gel Bond if needed to prevent it from extending beyond the glass plate. Material may also be trimmed to fit flush with the plate edge.

3. Assemble the gel cassette according to the manufacturer's instructions (also see Basic Protocol 1, steps 1 and 2). Just prior to assembly, blow air over the surface of both the Gel Bond and the opposing glass surface to remove any particulate material (e.g., dust).

   Sequencing gel spacers can be easily adapted. First, cut the spacers slightly longer

RCU IASH CENTRAL UNIVERSITY LIBRARY

than the length of the gel plate. Position a spacer along each edge of the glass plate and assemble the gel sandwich, clamping in place. With a razor blade, trim the excess spacer at top and bottom to get a reusable spacer exactly the size of the plate.

4. Prepare and pour the separating and stacking gels (see Basic Protocol 1, steps 3 to 9). In place of the Teflon comb, insert a square well sequencing comb cut to fit within the gel sandwich. Allow the stacking gel to polymerize 30 to 45 min at room temperature.

   Less solution is needed for ultrathin gels. For example, a 0.5-mm-thick gel requires 33% less gel solution than a 0.75-mm gel.

5. Prepare the sample and load the gel (see Basic Protocol 1, steps 12 to 20).

   When preparing protein samples for ultrathin gels, 3 to 4 µl at 5 µg protein/µl is required for Coomassie blue R-250 staining, whereas 10-fold less is needed for silver staining.

6. Run the gel (see Basic Protocol 1, steps 21 and 22), except conduct the electrophoresis at 7 mA/gel (0.25-mm-thick gels) or 14 mA/gel (0.5-mm-thick gels) for 4 to 5 hr.

7. When the separation is complete, disassemble the unit and remove the gel (see Basic Protocol 1, steps 23 to 26). With a gloved hand, wash away the adhesive material under a stream of water before proceeding to protein detection.

   Either Coomassie blue or silver staining (UNIT 10.5) may be used, but silver staining produces particularly fine resolution with thin Gel Bond–backed gels. Compared to staining thicker (>0.75 mm) gels, thin (<0.75 mm) gels stain and destain more quickly. Although the optimum staining times must be empirically determined, all steps in Coomassie blue and silver staining procedures are generally reduced by half.

   Gloves are needed throughout these procedures to prevent contamination by proteins on the surface of skin.

CASTING MULTIPLE SINGLE-CONCENTRATION GELS

SUPPORT PROTOCOL 1

Casting multiple gels at one time has several advantages. All the gels are identical, so sample separation is not affected by gel-to-gel variation. Furthermore, casting ten gels is only slightly more difficult than casting two gels. Once cast, gels can be stored for several days in a refrigerator.

Additional Materials (also see Basic Protocol 1)

Separating and stacking gels for single-concentration gels (Table 10.1.8)
H_2O-saturated isobutyl alcohol

Multiple gel caster (Bio-Rad, Hoefer Pharmacia)
Extra plates and spacers
14 × 14–cm acrylic blocks or polycarbonate sheets
250- and 500-ml side-arm flasks (used in gel preparation)
Long razor blade *or* plastic wedge (Wonder Wedge, Hoefer Pharmacia)
Resealable plastic bags

Pour the separating gel

1. Assemble the multiple gel caster according to the manufacturer's instructions.

   With the Hoefer Pharmacia unit make sure to insert the large triangular space filler plug in the base of the caster. The plug is removed when casting gradient gels (see Support Protocol 2).

2. Assemble glass sandwiches and stack them in the casting chamber. Stack up to ten

BCU ARTS CENTRAL UNIVERSITY LIBRARY

Table 10.1.8 Recipes for Multiple Single-Concentration Polyacrylamide Gels[a]

SEPARATING GEL

Stock solution[b]	Final acrylamide concentration in separating gel (%)[c]										
	5	6	7	8	9	10	11	12	13	14	15
30% acrylamide/ 0.8% bisacrylamide	52	62	72	83	93	103	114	124	134	145	155
4× Tris·Cl/SDS, pH 8.8	78	78	78	78	78	78	78	78	78	78	78
H_2O	181	171	160	150	140	129	119	109	98	88	78
10% (w/v) ammonium persulfate[d]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TEMED	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21

Preparation of separating gel

In a 500-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 4× Tris·Cl/SDS, pH 8.8 (Table 10.1.1), and H_2O. Degas under vacuum about 5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix; use immediately.

STACKING GEL

In a 250-ml side-arm flask, mix 13.0 ml 30% acrylamide/0.8% bisacrylamide solution, 25 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.1.1), and 61 ml H_2O. Degas under vacuum about 5 min. Add 0.25 ml 10% ammonium persulfate and 50 µl TEMED. Swirl gently to mix. Use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

[a]The recipes produce about 300 ml of separating gel and 100 ml of stacking gel, which are adequate for ten gels of dimensions 1.5 mm × 14 cm × 14 cm. Volumes were measured using 1.5-mm spacers (or fewer gels). For thinner spacers or fewer gels, calculate volumes using the equation in the annotation to step 4. The recipes are based on the SDS (denaturing) discontinuous buffer system of Laemmli (1970).

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in separating gel depends on the molecular size of the protein being separated. See Basic Protocol 1, annotation to step 3.

[d]Best to prepare fresh.

1.5-mm gels and fill in extra space with acrylic blocks or polycarbonate sheets to hold the sandwiches tightly in place. Make sure the spacers are straight along the top, right, and left edges of the glass plates and that all edges of the stack are flush.

The presence of loosely fitting sandwiches in the caster will lead to unevenly cast gels, creating distortions during electrophoresis. Polycarbonate inhibits gel polymerization. Therefore, if polycarbonate sheets are placed in the caster before and after the set of glass sandwiches, the entire set will slide out as one block after polymerization. Placing polycarbonate sheets between each gel sandwich makes them easier to separate from one another after polymerization.

3. Place the front sealing plate on the casting chamber, making sure the stack fits snugly. Secure the plate with four spring clamps and tighten the bottom thumb screws.

4. Prepare the separating (resolving) gel solution (Table 10.1.8).

A 12-cm separating gel with a 4-cm stacking gel is recommended.

If fewer than ten gels are prepared (Table 10.1.8), use the following formula to estimate the amount of separating gel volume needed: Volume = gel number × height (cm) × width (cm) × thickness (cm) + 4 × gel number + 10 ml.

PCU IASLIC CENTRAL UNIVERSITY LIBRARY

5. Using a 100-ml disposable syringe with flat-tipped needle, inject the resolving gel solution down the side of one spacer into the multiple caster. A channel in the silicone plug distributes the solution throughout the whole caster. Avoid introducing bubbles by giving the caster a quick tap on the benchtop once the caster is filled.
6. Overlay the center of each gel with 100 µl H_2O-saturated isobutyl alcohol and let polymerize for 1 to 2 hr.
7. Drain off the overlay and rinse the surface with 1× Tris·Cl/SDS, pH 8.8. If the gels will not be used immediately, skip to step 12.

Pour the stacking gel

8. Prepare the stacking gel solution either singly (see Basic Protocol 1, step 8) or for all the gels at once (Table 10.1.8).

   The stacking gel solution should be prepared just before pouring the gel.

9. Fill each sandwich in the caster with stacking gel solution.
10. Insert a comb into each sandwich and let the gel polymerize for 2 hr.

    Insert the combs at a 45° angle to avoid trapping air underneath the comb teeth. Air bubbles will inhibit polymerization and cause dents in the wells and a distorted pattern of protein bands.

11. Remove the combs and rinse wells with 1× SDS electrophoresis buffer.

Remove the gels from the caster

12. Remove the gels from the caster and separate by carefully inserting a long razor blade or knife between each gel sandwich. A plastic wedge (Hoefer Pharmacia's Wonder Wedge) also works well.
13. Clean the outside of each gel plate with running water to remove the residual polymerized and unpolymerized acrylamide.
14. Overlay gels to be stored with 1× Tris·Cl/SDS, pH 8.8, place in a resealable plastic bag, and store at 4°C until needed (up to 1 week).

SEPARATION OF PROTEINS ON GRADIENT GELS

ALTERNATE PROTOCOL 5

Gels that consist of a gradient of increasing polyacrylamide concentration resolve a much wider size range of proteins than standard uniform concentration gels (see Critical Parameters and Troubleshooting). The protein bands, particularly in the low-molecular-weight range, are also much sharper. Unlike single-concentration gels, gradient gels separate proteins in a way that can be plotted easily to give a linear curve from 10 to 200 kDa. This facilitates molecular weight estimations.

The quantities given below provide separating gel solution sufficient for two 0.75-mm gels (~7 ml of each concentration) or one 1.5-mm gel (~14 ml of each concentration). Volumes can be adjusted to accommodate gel sandwiches of different dimensions.

Additional Materials (also see Basic Protocol 1)

Light and heavy acrylamide gel solutions (Table 10.1.9 and Table 10.1.10)
Bromphenol blue (optional; for checking practice gradient)
TEMED

Gradient maker (30 to 50 ml, Hoefer Pharmacia SG30 or SG50; *or* 30 to 100 ml, Bio-Rad 385)
Tygon tubing with micropipet tip

BCU PSI CENTRAL UNIVERSITY LIBRARY

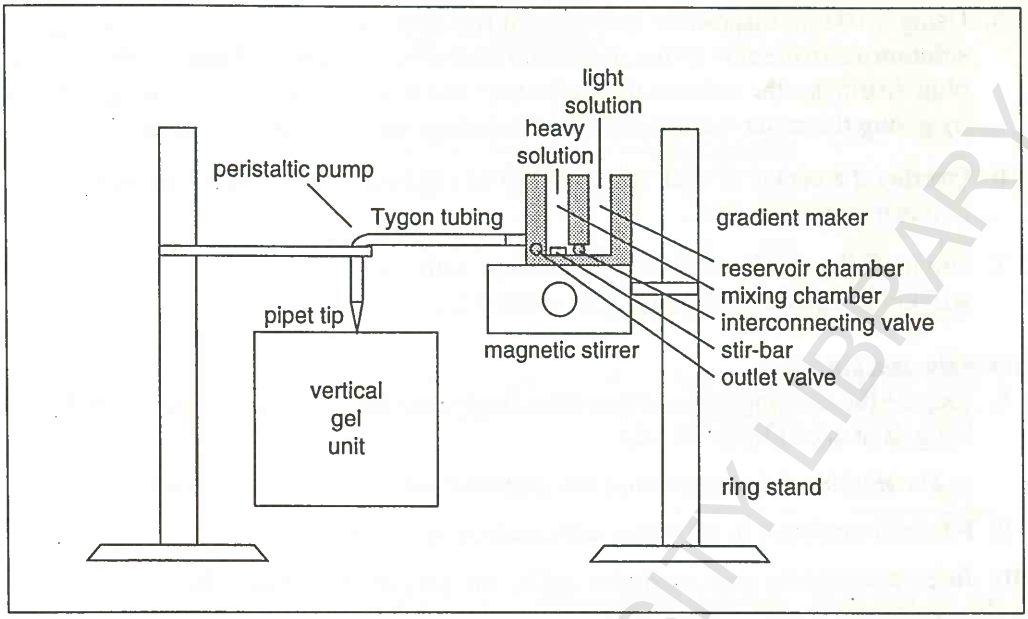


Figure 10.1.2 Gradient gel setup. A peristaltic pump, though not required, will provide better control.

Peristaltic pump (optional; Markson A-13002, A-34040, or A-34105 minipump)
Whatman 3MM filter paper

Set up the gradient maker and prepare the gel solutions

1. Assemble the magnetic stirrer and gradient maker on a ring stand as shown in Figure 10.1.2. Connect the outlet valve of the gradient maker to Tygon tubing attached to a micropipet tip that is placed over the vertical gel sandwich. If desired, place a peristaltic pump in line between the gradient maker and the gel sandwich.

   A peristaltic pump will simplify casting by providing a smooth flow rate.

2. Place a small stir-bar into the mixing chamber of the gradient maker (i.e., the chamber connected to the outlet).

3. Using the recipes in Table 10.1.9 and Table 10.1.10, prepare light and heavy acrylamide gel solutions. Do not add ammonium persulfate until just before use (step 7).

4. With the outlet port and interconnecting valve between the two chambers closed, pipet 7 ml of light (low-concentration) acrylamide gel solution into the reservoir chamber for one 0.75-mm-thick gradient gel.

   Recommended gradient ranges are 5% to 20% for most applications (to separate proteins of 10 to several hundred kilodaltons).

   A practice run with heavy and light solutions is recommended. Bromphenol blue should be added to the heavy solution to demonstrate linearity of the practice gradient.

5. Open the interconnecting valve briefly to allow a small amount (~200 µl) of light solution to flow through the valve and into the mixing chamber.

   The presence of air bubbles in the interconnecting valve may obstruct the flow between chambers during casting.

   Deaeration is not recommended for either the light or heavy solution. Omitting the deaeration will allow polymerization to proceed more slowly, letting the gradient establish itself in the gel sandwich before polymerization takes place.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 10.1.9 Light Acrylamide Gel Solutions for Gradient Gels[a]

Stock solution	Acrylamide concentration of light gel solution (%)[b]									
	5	6	7	8	9	10	11	12	13	14
30% acrylamide/ 0.8% bisacrylamide[c]	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
4× Tris·Cl/SDS, pH 8.8[c]	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H_2O	8.75	8.25	7.75	7.25	6.75	6.25	5.75	5.25	4.75	4.25
10% (w/v) ammonium persulfate	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

[a]To survey proteins ≥10 kDa, 5-20% gradient gels are recommended. To expand the range between 10 and 200 kD, a 10-20% gel is recommended.

[b]Numbers in body of table are milliliters of stock solution. Deaeration is not required. Keep solution at room temperature prior to adding TEMED no longer than 1 hr.

[c]See Table 10.1.1 for preparation.

Table 10.1.10 Heavy Acrylamide Gel Solutions for Gradient Gels[a]

Stock solution	Acrylamide concentration of heavy gel solution (%)[b]										
	10	11	12	13	14	15	16	17	18	19	20
30% acrylamide/ 0.8% bisacrylamide[c]	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
4× Tris·Cl/SDS, pH 8.8[c]	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H_2O	5.0	4.5	4.0	3.5	3.0	2.5	2.0	1.5	1.0	0.5	0
Sucrose (g)	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25
10% (w/v) ammonium persulfate[b]	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

[a]Deaeration is not recommended for gradient gels.

[b]Numbers in body of table are milliliters of stock solution (except sucrose). Do not add the ammonium persulfate until just before use. The heavy acrylamide will polymerize, albeit more slowly, without the addition of TEMED. Keep the heavy solution on ice after adding ammonium persulfate.

[c]See Table 10.1.1 for preparation.

6. Add 7 ml of heavy (high-concentration) acrylamide gel solution to the mixing chamber.

   Keep the heavy solution on ice until use. Once the ammonium persulfate is added to the heavy solution, it will polymerize without TEMED, albeit more slowly; keeping the solution on ice prevents this. The gel solution will come to room temperature during casting. The higher the percentage of acrylamide, the more severe the problem of premature polymerization is.

7. Add specified amount of ammonium persulfate and ~2.3 µl TEMED per 7 ml acrylamide solution to each chamber. Mix the solutions in each chamber with a disposable pipet.

Form the gradient and cast the gel

8. Open the interconnecting valve completely.

   Some of the heavy solution will flow back into the reservoir chamber containing light solution as the two chambers equilibrate. This will not affect the formation of the gradient.

BOJAN CENTRAL UNIVERSITY LIBRARY

9. Turn on the magnetic stirrer and adjust the rate to produce a slight vortex in the mixing chamber.

10. Open the outlet of the gradient maker slowly. Adjust the outlet valve to a flow rate of 2 ml/min. If using a peristaltic pump, calibrate the flow rate with a graduated cylinder prior to casting the gel.

    Some adjustment of the flow rate may be necessary during casting. If the light solution is not flowing into the mixing chamber, a bubble may be caught in the interconnecting valve. Quickly close the outlet and cover the top of the reservoir chamber with a gloved thumb. Push down with the thumb to increase the pressure in the chamber and force the air bubble out of the center valve.

11. Fill the gel sandwich from the top. Place the pipet tip against one side of the sandwich so the solution flows down one plate only. The heavy solution will flow into the sandwich first, followed by progressively lighter solution.

12. Watch as the last of the light solution drains into the outlet tube and adjust the flow rate to ensure that the last few milliliters of solution do not flow quickly into the gel sandwich and disturb the gradient.

13. Overlay the gradient gel with H_2O-saturated isobutyl alcohol. Allow the gel to polymerize ~1 hr.

    In this gel system, the gel will polymerize from the bottom (i.e., heavy solution) up. Because polymerization is an exothermic reaction, heat can be felt evolving from the bottom of the gel sandwich during polymerization. A sharp optical discontinuity at the gel-overlay interface indicates that polymerization has occurred. In general, 1 hr is adequate for polymerization.

14. Remove the H_2O-saturated isobutyl alcohol and rinse with 1× Tris·Cl/SDS, pH 8.8. Cast the stacking gel (see Basic Protocol 1, steps 8 to 11).

    The gel can be covered with 1× Tris·Cl/SDS, pH 8.8, sealed in a plastic bag, and stored for up to 1 week.

Load and run the gel

15. Prepare the protein sample and protein molecular-weight-standards mixture (see Basic Protocol 1, steps 13 to 26). Load and run the gel.

    The gel can be stained with Coomassie blue or silver (UNIT 10.5).

16. After staining, dry the gels onto Whatman 3MM or equivalent filter paper.

    Gradient gels >0.75 mm thick require special handling during drying to prevent cracking. The simplest approach to drying gradient gels is to use thin gels; ≤0.75-mm gradient gels with ≤20% acrylamide solutions will dry without cracking as long as the vacuum pump is working properly and the cold trap is dry at the onset of drying. For gradient gels >0.75 mm thick, add 3% (w/v) glycerol to the final destaining solution to help prevent cracking. Another method is to dehydrate and shrink the gel in 30% methanol for up to 3 hr prior to drying. Then place the gel in distilled (Milli-Q-purified or equivalent) water for 5 min before drying.

PGU IASI / CENTRAL UNIVERSITY LIBRARY

CASTING MULTIPLE GRADIENT GELS

Casting gradient gels in a multiple gel caster has several advantages. In addition to the time savings, batch casting gives gels that are essentially identical. This is particularly important for gradient gels, where slight variations in casting technique can cause variations in protein mobility. The gels may be stored for up to 1 week after casting to ensure internal consistency from run to run during the week. Furthermore, gels with several ranges of concentrations (e.g., 5% to 20% and 10% to 20% acrylamide) can be cast and stored, giving much more flexibility to optimize separations.

Additional Materials (also see Alternate Protocol 5)

Plug solution (see recipe)
Light and heavy acrylamide gel solutions for multiple gradient gels (Table 10.1.11 and Table 10.1.12)
TEMED
H_2O-saturated isobutyl alcohol

Multiple gel caster (Bio-Rad, Hoefer Pharmacia)
Peristaltic pump (25 ml/min)
500- or 1000-ml gradient maker (Bio-Rad, Hoefer Pharmacia)
Tygon tubing

Set up system and pour separating gel

1. Assemble the multiple caster as in casting multiple single-concentration gels (see Support Protocol 1, steps 1 to 3), making sure to remove the triangular space filler plugs in the bottom of the caster.

   The plug is used only when casting single-concentration gels.

2. Set up the peristaltic pump (Fig. 10.1.3). Using a graduated cylinder and water, adjust the flow rate so that the volume of the gradient solution plus volume of plug solution is poured in ~15 to 18 min (~25 ml/min).

3. Set up the gradient maker. Close all valves and place a stir-bar in the mixing chamber, which is the one with the outlet port. Attach one end of a piece of Tygon tubing to the outlet of the gradient maker. Run the other end of the tubing through the peristaltic pump and attach it to the red inlet port at the bottom of the caster.

   Choose a gradient maker that holds no more than four times the total volume of the gradient solution to be poured (i.e., a 1000- or 500-ml gradient maker).

4. Prepare solutions for the gradient maker (Table 10.1.11 and Table 10.1.12).

5. Add the TEMED to both heavy and light solutions (54 µl/165 ml) and immediately pour the light (low concentration) solution into the mixing chamber (the one with the port). Open the mixing valve slightly to allow the tunnel to fill and to avoid air bubbles. Close the valve again and pour the heavy (high-concentration) acrylamide solution into the reservoir chamber.

6. Start the magnetic stirrer and open the outlet valve; then start the pump and open the mixing valve.

   In units for casting multiple gels, acrylamide solution flows in from the bottom. To use a multiple casting unit, the light solution is placed in the mixing chamber and the heavy in the reservoir. This is the reverse of casting a single gel (see Alternate Protocol 5). Thus, light solution enters the multiple caster first, followed by progressively heavier solution. Finally, the acrylamide solution is stabilized in the multiple caster with a

BCU INSTITUTE CENTRAL UNIVERSITY LIBRARY

Table 10.1.11 Light Acrylamide Gel Solutions for Multiple Gradient Gels[a,b]

Stock solution	Acrylamide concentration of light separating gel solution (%)[c]									
	5	6	7	8	9	10	11	12	13	14
30% acrylamide/0.8% bisacrylamide[d]	28	33	39	44	50	55	61	66	72	77
4× Tris·Cl/SDS, pH 8.8[d]	41	41	41	41	41	41	41	41	41	41
H_2O	96	91	85	80	74	69	63	58	52	47
10% (w/v) ammonium persulfate	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55

[a]To survey proteins ≥10 kDa, 5-20% gradient gels are recommended. To expand the range between 10 and 200 kDa, a 10-20% gel is recommended.

[b]Recipes produce ten 1.5-mm-thick gradient gels with 10 ml extra solution to account for losses in tubing.

[c]Numbers in body of table are milliliters of stock solution. Deaeration is not required. Keep solution at room temperature prior to adding TEMED no longer than 1 hr.

[d]See Table 10.1.1 for preparation.

heavy plug solution and allowed to polymerize (see step 8 and manufacturer's instructions).

7. When almost all the acrylamide solution is gone from the gradient maker, stop the pump and close the mixing valve. Tilt the gradient maker toward the outlet side and remove the last milliliters of the mix. Do not allow air bubbles to enter the tubing.

8. Add the plug solution to the mixing chamber and start the pump. Make sure that no bubbles are introduced. Continue pumping until the bottom of the caster is filled with plug solution to just below the glass plates; then turn off the pump. Clamp the tubing close to the red port of the casting chamber.

9. Quickly overlay each separate gel sandwich with 100 µl H_2O-saturated isobutyl alcohol. Use the same amount on each sandwich.

   Failure to use the same amount of overlay solution will cause the gel sandwiches to polymerize at different heights.

10. Drain off the overlay and rinse the surface of the gels with 1× Tris·Cl/SDS, pH 8.8.

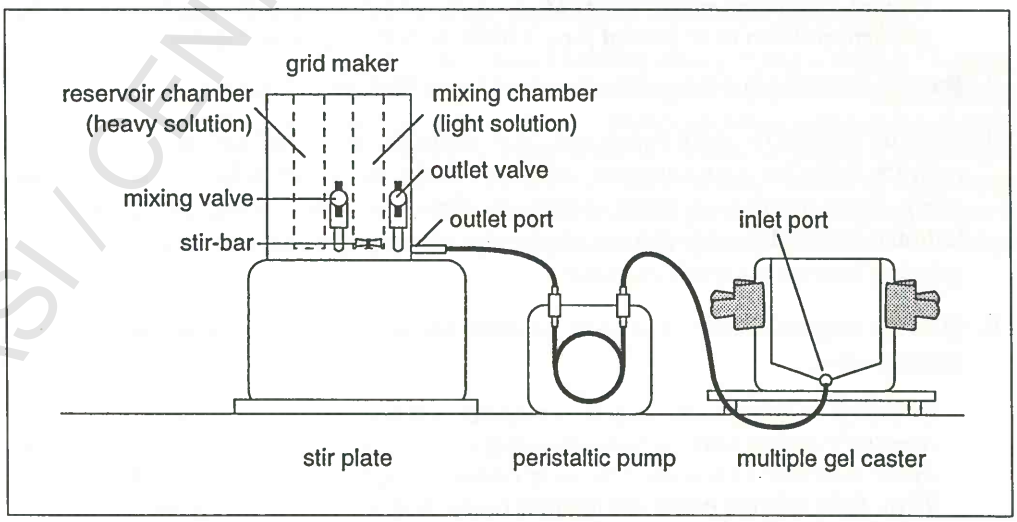


Figure 10.1.3 Setup for casting multiple gradient gels. Casting multiple gradient gels requires a peristaltic pump and a multiple gel caster. Gel solution is introduced through the bottom of the multiple caster.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 10.1.12 Heavy Acrylamide Gel Solutions for Multiple Gradient Gels[a,b]

Stock solution	Acrylamide concentration of heavy gel solution (%)[c]										
	10	11	12	13	14	15	16	17	18	19	20
30% acrylamide/0.8% bisacrylamide[d]	55	61	66	72	77	83	88	94	99	105	110
4× Tris·Cl/SDS pH 8.8[d]	41	41	41	41	41	41	41	41	41	41	41
H_2O	55	50	44	39	33	28	22	17	11	5.5	0
Sucrose (g)	25	25	25	25	25	25	25	25	25	25	25
10% (w/v) ammonium persulfate	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55

[a]Deaeration is not recommended for gradient gels.

[b]Recipes produce 10 ml extra solution to account for losses in tubing.

[c]Numbers in body of table are milliliters of stock solution (except sucrose). Do not add the ammonium persulfate until just before use. The heavy acrylamide will polymerize, albeit more slowly, without the addition of TEMED. Keep the heavy solution on ice after adding ammonium persulfate.

[d]See Table 10.1.1 for preparation.

Pour stacking gel and remove gels from caster

11. Prepare and cast the stacking gel as in casting multiple single-concentration gels (see Support Protocol 1, steps 8 to 11).

12. Remove gels from the caster and clean the gel sandwiches (see Support Protocol 1, steps 12 and 13). Store gels, if necessary, according to the instructions for multiple single-concentration gels (see Support Protocol 1, step 14).

ELECTROPHORESIS IN SINGLE-CONCENTRATION MINIGELS

BASIC PROTOCOL 2

Separation of proteins in a small-gel format is becoming increasingly popular for applications that range from isolating material for peptide sequencing to performing routine protein separations. The unique combination of speed and high resolution is the foremost advantage of small gels. Additionally, small gels are easily adapted to single-concentration, gradient, and two-dimensional SDS-PAGE procedures. The minigel procedures described are adaptations of larger gel systems.

This protocol describes the use of a multiple gel caster. The caster is simple to use, and up to five gels can be prepared at one time with this procedure. Single gels can be prepared using adaptations in the manufacturer's instructions. A multiple gel caster is the only practical way to produce small linear polyacrylamide gradient gels (see Support Protocol 3).

Materials

Minigel vertical gel unit (Hoefer Pharmacia Mighty Small SE 250/280 *or* Bio-Rad Mini-Protean II) with glass plates, clamps, and buffer chambers
0.75-mm spacers
Multiple gel caster (Hoefer Pharmacia SE-275/295 *or* Bio-Rad Mini-Protean II multicasting chamber)
Acrylic plate (Hoefer Pharmacia SE-217 *or* Bio-Rad 165-1957) or polycarbonate separation sheet (Hoefer Pharmacia SE-213 *or* Bio-Rad 165-1958)
10- and 50-ml syringes
Combs (Teflon, Hoefer Pharmacia SE-211A series or Bio-Rad Mini-Protean II)
Long razor blade
Micropipet

BOYNTON / CENTRAL UNIVERSITY LIBRARY

Additional reagents and equipment for standard denaturing SDS-PAGE (see Basic Protocol 1)

Pour the separating gel

1. Assemble each gel sandwich by stacking, in order, the notched (Hoefer Pharmacia) or small rectangular (Bio-Rad) plate, 0.75-mm spacers, and the larger rectangular plate. Be sure to align the spacers properly with the ends flush with the top and bottom edge of the two plates when positioning the sandwiches in the multiple gel caster (Fig. 10.1.4).

   The protocol described is basically for the Hoefer Pharmacia system. For other systems, make adjustments according to the manufacturer's instructions. Alternatively, precast minigels can be purchased from a number of suppliers (see Table 10.1.3).

   The multiple casters from Hoefer Pharmacia have a notch in the base designed for casting gradient gels. A silicone rubber insert fills up this space when casting single-concentration gels. The Hoefer Pharmacia spacers are T-shaped to prevent slipping. The flanged edge of the spacer must be positioned against the outside edge of the glass plate. Placing a sheet of wax paper between the gel sandwiches will help separate the sandwiches after polymerization.

2. Fit the gel sandwiches tightly in the multiple gel caster. Use an acrylic plate or polycarbonate separation sheet to eliminate any slack in the chamber.

   Loosely fitting sandwiches in the caster will lead to unevenly cast gels, creating distortions during electrophoresis.

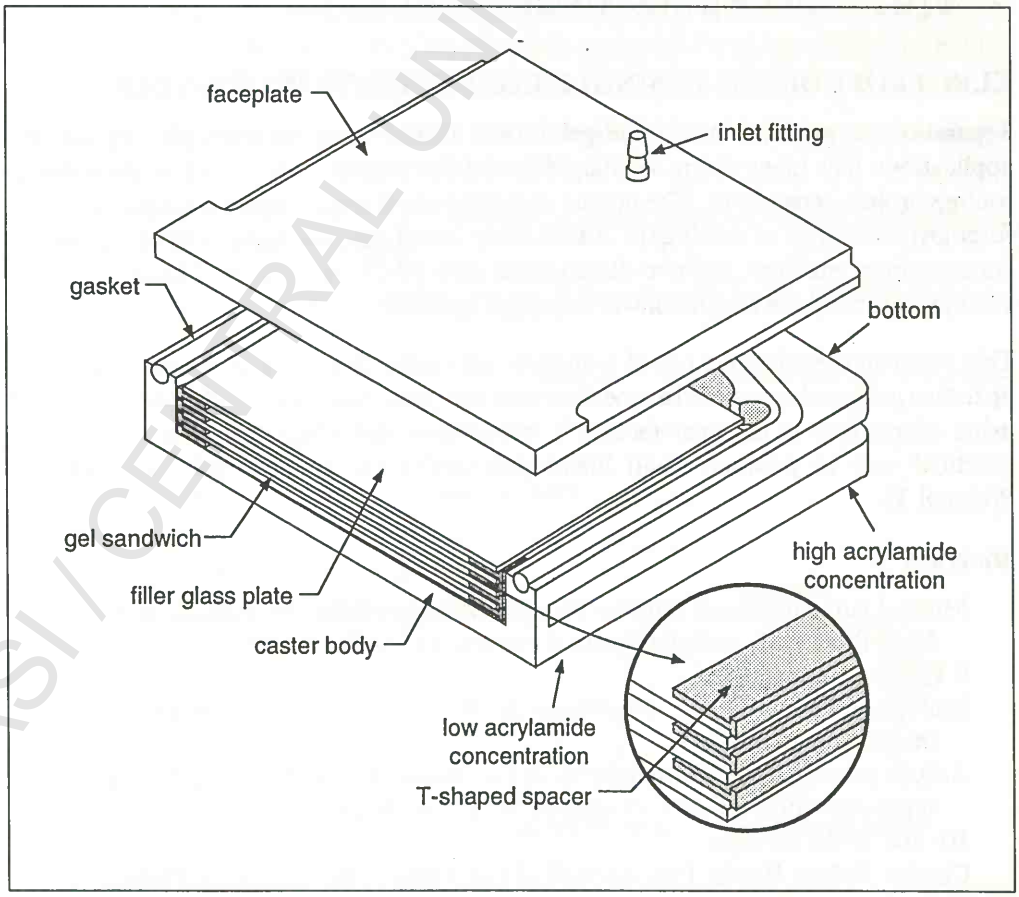


Figure 10.1.4 Minigel sandwiches positioned in the multiple gel caster. Extra glass or acrylic plates or polycarbonate sheets are used to fill any free space in the caster and to ensure that the gel sandwiches are held firmly in place.

ACU IASI / CENTRAL UNIVERSITY LIBRARY

3. Place the front faceplate on the caster, clamp it in place against the silicone gasket, and verify alignment of the glass plates and spacers.

4. Prepare the separating gel solution as directed in Table 10.1.1. For five 0.75-mm-thick gels, prepare ~30 ml solution (i.e., double the volumes listed).

   To compute the total gel volume needed, multiply the area of the gel (e.g., 7.3 × 8.3 cm) by the thickness of the gel (e.g., 0.75 mm) and then by the number of gels in the caster. If needed, add about 4 to 5 ml of extra gel solution to account for the space around the outside of the gel sandwiches.

   Do not add TEMED and ammonium persulfate until just before use.

5. Fill a 50-ml syringe with the separating gel solution and slowly inject it into the caster until the gels are 6 cm high, allowing 1.5 cm for the stacking gel.

6. Overlay each gel with 100 µl H_2O-saturated isobutyl alcohol. Allow the gels to polymerize for ~1 hr.

Pour the stacking gel

7. Remove the isobutyl alcohol and rinse with 1× Tris·Cl/SDS, pH 8.8.

   Stacking gels can be cast one at a time with the gel mounted on the electrophoresis unit, or all at once in the multiple caster.

8. Practice placing a comb in the gel sandwiches before preparing the stacking gel solution. Press the comb against the rectangular or taller plate so that all teeth of the comb are aligned with the opening in the gel sandwich, then insert into the sandwich. Remove combs after practicing.

9. Prepare the stacking gel solution (2 ml per gel) as directed in Table 10.1.1. Fill a 10-ml syringe with stacking gel solution and inject the solution into each gel sandwich.

10. Insert combs, taking care not to trap bubbles. Allow gels to polymerize 1 hr.

11. Remove the front faceplate. Carefully pull the gels out of the caster, using a long razor blade to separate the sandwiches.

    If the gels are left to polymerize for prolonged periods, they will be difficult to remove from the caster.

    The gels can be stored tightly wrapped in plastic wrap with the combs left in place inside a sealable bag to prevent drying for ~1 week. Without the stacking gel, the separating gel can be stored for 2 to 3 weeks. Keep gels moist with 1× Tris·Cl/SDS, pH 8.8, at 4°C. Do not store gels in the multiple caster.

Prepare the sample, load the gel, and conduct electrophoresis

12. Remove the combs and rinse the sample wells with 1× SDS electrophoresis buffer. Place a line indicating the bottom of each well on the front glass plate with a marker.

13. Fill the upper and lower buffer chambers with 1× SDS electrophoresis buffer. The upper chamber should be filled to 1 to 2 cm over the notched plate.

14. Prepare the protein sample and protein-standards mixture (see Basic Protocol 1, step 12).

15. Load the sample using a micropipet. Insert the pipet tip through the upper buffer and into the well. The mark on the glass plate will act as a guide. Dispense the sample into the well.

    For a complex mixture, 20 to 25 µg protein in 10 µl SDS sample buffer will give a strongly stained Coomassie blue pattern. Much smaller amounts (1 to 5 µg) are required for highly purified proteins, and 10- to 100-fold less protein amount in the same volume (e.g., 10 µl is required for silver staining).

BOCAS RATON CENTRAL UNIVERSITY LIBRARY

16. Electrophorese samples at 10 to 15 mA per 0.75-mm gel until the dye front reaches the bottom of the gel (~1 to 1.5 hr).

17. Disassemble the gel (see Basic Protocol 1, steps 23 to 26). Proceed with detection of proteins.

SUPPORT PROTOCOL 3

PREPARING MULTIPLE GRADIENT MINIGELS

Polyacrylamide gradients not only enhance the resolution of larger format gels but also greatly improve protein separation in the small format. Casting gradient minigels one at a time is not generally feasible because of the small volumes used, but multiple gel casters make it easy to cast several small gradient gels at one time. The gels are cast from the bottom in multiple casters, with the light acrylamide solution entering first. This is the opposite of casting one gel at a time, in which the heavy solution enters from the top of the gel sandwich and flows down to the bottom.

Additional Materials (also see Basic Protocol 2)

Plug solution (see recipe)

Additional reagents and equipment for preparing gradient gels (see Alternate Protocol 5)

Set up the system and prepare the gel solutions

1. Assemble minigel sandwiches in the multiple gel caster as described for single-concentration minigels (see Basic Protocol 2, steps 1 to 3).

2. Set up the 30-ml gradient maker, magnetic stirrer, peristaltic pump (optional), and Tygon tubing as in Figure 10.1.3. Connect the outlet of the 30-ml gradient maker to the inlet at the base of the front faceplate of the caster.

   The monomer solution will be introduced through the inlet at the bottom of the front faceplate of the caster first, followed by progressively heavier solution.

3. Prepare light (Table 10.1.9) and heavy (Table 10.1.10) acrylamide gel solutions. Use ~12 ml of each solution for five 0.75-mm-thick minigels.

   Adjust volumes if a different thickness or number of gels is needed. Do not add ammonium persulfate until just before use. Deaeration is not recommended for gradient gels.

4. With the outlet and interconnecting valve closed, add the heavy solution to the reservoir chamber. Briefly open the interconnecting valve to let a small amount of heavy solution through to the mixing chamber, clearing the valve of air.

5. Fill the mixing chamber with light solution. Add 4 µl TEMED per 12 ml acrylamide solution to each chamber and mix with a disposable pipet.

Form the gradient and cast the gels

6. Turn on the magnetic stirrer. Open the interconnecting valve and allow the chambers to equilibrate. Then slowly open the outlet port to allow the solution to flow from the gradient maker to the multiple caster by gravity (a peristaltic pump may be used for better control). Adjust the flow rate to 3 to 4 ml/min.

   Faster flow rates are possible and will also produce good gradients. However, a fast flow increases the potential for introduction of bubbles into the caster.

7. Close the outlet port as the last of the gradient solution leaves the mixing chamber, just before air enters the outlet tube. Fill the two chambers with plug solution and slowly open the outlet once again.

PGU IAS / CENTRAL UNIVERSITY LIBRARY

8. Allow the plug solution to push the acrylamide in the caster up into the plates. Close the outlet when the plug solution reaches the bottom of the plates.

   A discontinuity between the bottom of the gels and the plug solution will be obvious.

9. Quickly add 100 µl H_2O-saturated isobutyl alcohol to each gel sandwich. Let the gels polymerize undisturbed for ~1 hr.

10. Prepare and pour the stacking gel (see Basic Protocol 2, steps 9 and 10).

Disassemble the system

11. Disconnect the gradient maker, place the caster in a sink, and remove the front faceplate. The plug solution will drain out from bottom of the caster.

12. Remove the gels (see Basic Protocol 2, step 11).

    Gradient minigels can be stored as described for single-concentration minigels (see Basic Protocol 2, step 11 annotation). For instructions on preparing, loading, and running the gels, see Basic Protocol 2, steps 12 to 17.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Anode buffer

121.1 g Tris base
500 ml H_2O
Adjust to pH 8.9 with concentrated HCl
Dilute to 5 liters with H_2O
Store at 4°C up to 1 month

Final concentration is 0.2 M Tris·Cl, pH 8.9.

Cathode buffer

12.11 g Tris base
17.92 g tricine
1 g SDS
Dilute to 1 liter with H_2O
Do not adjust pH
Store at 4°C up to 1 month

Final concentrations are 0.1 M Tris, 0.1 M tricine, and 0.1% (w/v) SDS.

Coomassie blue G-250 staining solution

200 ml acetic acid
1800 ml H_2O
0.5 g Coomassie blue G-250
Mix 1 hr and filter (Whatman no. 1 paper)
Store at room temperature indefinitely

The final solution is 0.025% (w/v) Coomassie blue G-250 in 10% (v/v) acetic acid.

Phosphate/SDS electrophoresis buffer

Dilute 500 ml of 4× phosphate/SDS, pH 7.2 (Table 10.1.7) with H_2O to 2 liters. Store at 4°C up to 1 month.

Final concentrations are 0.1 sodium phosphate (pH 7.2)/0.1% (w/v) SDS.

BOU IAST / CENTRAL UNIVERSITY LIBRARY

Phosphate/SDS sample buffer, 2× (for continuous systems)

0.5 ml 4× phosphate/SDS, pH 7.2 (Table 10.1.7)
0.2 g SDS
0.1 mg bromphenol blue
0.31 g DTT
2.0 ml glycerol
Add H_2O to 10 ml and mix

Final concentrations are 20 mM sodium phosphate, 2% (w/v) SDS, 0.001% (w/v) bromphenol blue, 0.2 M DTT, and 2% (v/v) glycerol.

Plug solution

0.125 M Tris·Cl, pH 8.8 (*APPENDIX 2E*)
50% (w/v) sucrose
0.001% (w/v) bromphenol blue
Store at 4°C up to 1 month

Recrystallized SDS (optional)

High-purity SDS is available from several suppliers, but for some sensitive applications (e.g., protein sequencing) recrystallization is useful. Commercially available electrophoresis-grade SDS is usually of sufficient purity for most applications.

Add 100 g SDS to 450 ml ethanol and heat to 55°C. While stirring, gradually add 50 to 75 ml hot H_2O until all SDS dissolves. Add 10 g activated charcoal (Norit 1, Sigma) to solution. After 10 min, filter solution through Whatman no. 5 paper on a Buchner funnel to remove charcoal. Chill filtrate 24 hr at 4°C and 24 hr at –20°C. Collect crystalline SDS on a coarse-frit (porosity A) sintered-glass funnel and wash with 800 ml –20°C ethanol (reagent grade). Repeat crystallization without adding activated charcoal. Dry recrystallized SDS under vacuum overnight at room temperature. Store in a desiccator over P_2O_5 in a dark bottle.

If proteins will be electroeluted or electroblotted for protein sequence analysis, it may be desirable to crystallize the SDS twice from ethanol/H_2O (Hunkapiller et al., 1983).

SDS electrophoresis buffer, 5×

15.1 g Tris base
72.0 g glycine
5.0 g SDS
H_2O to 1000 ml
Dilute to 1× or 2× for working solution, as appropriate

Do not adjust the pH of the stock solution, as the solution is pH 8.3 when diluted. Store at 0° to 4°C until use (up to 1 month).

SDS sample buffer, 2× (for discontinuous systems)

25 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.1.1)
20 ml glycerol
4 g SDS
2 ml 2-ME *or* 3.1 g DTT
1 mg bromphenol blue
Add H_2O to 100 ml and mix
Store in 1-ml aliquots at –70°C

To avoid reducing proteins to subunits (if desired), omit 2-ME or DTT (reducing agent) and add 10 mM iodoacetamide to prevent disulfide interchange.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

SDS sample buffer, 6× (for discontinuous systems)

7 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.1.1)
3.0 ml glycerol
1 g SDS
0.93 g DTT
1.2 mg bromphenol blue
Add H_2O to 10 ml (if needed)
Store in 0.5-ml aliquots at –70°C

Tricine sample buffer, 2×

2 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.1.1)
2.4 ml (3.0 g) glycerol
0.8 g SDS
0.31 g DTT
2 mg Coomassie blue G-250
Add H_2O to 10 ml and mix

Final concentrations are 0.1 M Tris, 24% (w/v) glycerol, 8% (w/v) SDS, 0.2 M DTT, and 0.02% (w/v) Coomassie blue G-250.

COMMENTARY

Background Information

Polyacrylamide gels form after polymerization of monomeric acrylamide into polymeric polyacrylamide chains and cross-linking of the chains by *N,N'*-methylenebisacrylamide. The polymerization reaction is initiated by the addition of ammonium persulfate, and the reaction is accelerated by TEMED, which catalyzes the formation of free radicals from ammonium persulfate. Because oxygen inhibits the polymerization process, deaerating the gel solution before the polymerization catalysts are added will speed up polymerization; deaeration is not recommended for the gradient gel protocols because slower polymerization facilitates casting of gradient gels.

Precast gels for commonly used vertical minigel and standard-sized SDS-PAGE apparatuses are available from several manufacturers (Table 10.1.3). Flatbed (horizontal) isoelectric focusing (IEF) and SDS-PAGE gels are not listed. Pharmacia supplies a range of horizontal gels for a variety of applications and should be consulted for further information. When using precast gels, pay strict attention to shelf life. In general, manufacturers overrate the shelf life, and the sooner the gels are used, the better. When reasonably fresh, precast gels provide excellent resolution that is as good or better than a typical gel cast in the laboratory.

The most widely used method for discontinuous gel electrophoresis is the system described by Laemmli (1970). This is the denaturing (SDS) discontinuous method used in Basic Protocol 1. A discontinuous buffer system uses buffers of different pH and composition to generate a discontinuous pH and voltage gradient in the gel. Because the discontinuous gel system concentrates the proteins in each sample into narrow bands, the applied sample may be more dilute than that used for continuous electrophoresis.

In the discontinuous system the sample first passes through a stacking gel, which has large pores. The stacking gel buffer contains chloride ions (called the leading ions) whose electrophoretic mobility is greater than the mobility of the proteins in the sample. The electrophoresis buffer contains glycine ions (called the trailing ions) whose electrophoretic mobility is less than the mobility of the proteins in the sample. The net result is that the faster migrating ions leave a zone of lower conductivity between themselves and the migrating protein. The higher voltage gradient in this zone allows the proteins to move faster and to "stack" in the zone between the leading and trailing ions. After leaving the stacking gel, the protein enters the separating gel. The separating gel has a smaller pore size, a higher salt concentration, and higher pH compared to the stacking gel. In the separating gel, the glycine ions migrate past the proteins, and the proteins are separated according to either molecular size in a denaturing gel (containing SDS) or molecular shape, size, and charge in a nondenaturing gel.

Proteins are denatured by heating in the presence of a low-molecular-weight thiol (2-ME or DTT) and SDS. Most proteins bind SDS in a constant-weight ratio, leading to identical

BOSTON UNIVERSITY LIBRARY

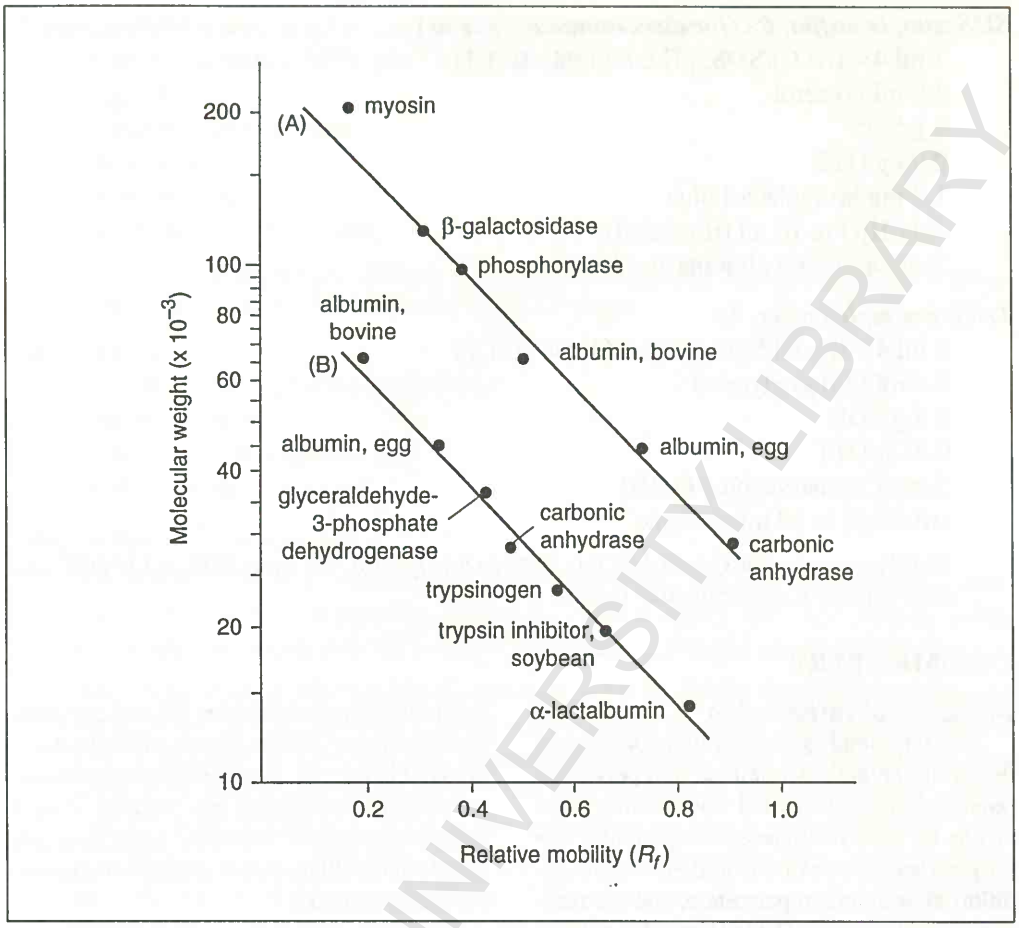


Figure 10.1.5 Typical calibration curves obtained with standard proteins separated by nongradient denaturing (SDS) discontinuous gel electrophoresis based on the method of Laemmli (1970). **(A)** Gel with 7% polyacrylamide. **(B)** Gel with 11% polyacrylamide. (Redrawn with permission from Sigma.)

charge densities for the denatured proteins. Thus, the SDS-protein complexes migrate in the polyacrylamide gel according to size, not charge. Most proteins are resolved on polyacrylamide gels containing from 5% to 15% acrylamide and 0.2% to 0.5% bisacrylamide (see Table 10.1.1). The relationship between the relative mobility and log molecular weight is linear over these ranges (Fig. 10.1.5). With the use of plots like those shown in Figure 10.1.5, the molecular weight of an unknown protein (or its subunits) may be determined by comparison with known protein standards (Table 10.1.2).

Basic Protocol 1 relies on denaturing proteins in the presence of SDS and 2-ME or DTT. Under these conditions, the subunits of proteins are dissociated and their biological activities are lost. A true estimate of a protein's molecular size can be made by comparing the relative mobility of the unknown protein to proteins in a calibration mixture (Fig. 10.1.5). Gradient gels (Alternate Protocol 5) simplify molecular-weight determinations by producing a linear relationship between log molecular weight of the protein and log % T over a much wider size range than single-concentration gels. Although percent acrylamide monomer is a more common measure of gel concentration, % T, the percentage of total monomer (acrylamide plus bisacrylamide) in the solution or gel, is used for molecular weight calculations in gradient gels. The % T is estimated assuming the acrylamide gradient is linear. For example, proteins in the gel shown in Figure 10.1.6 were separated in a 5% to 20% T acrylamide gradient. The % T of the point halfway through the resolving gel is 12.5% T. Simply plotting log molecular mass versus distance moved into the gel (or R_f) also produces a relatively linear standard curve over a fairly wide size range.

If two proteins have identical molecular sizes, they more than likely will not be resolved with one-dimensional SDS-PAGE, and two-di-

PCU IASRI CENTRAL UNIVERSITY LIBRARY

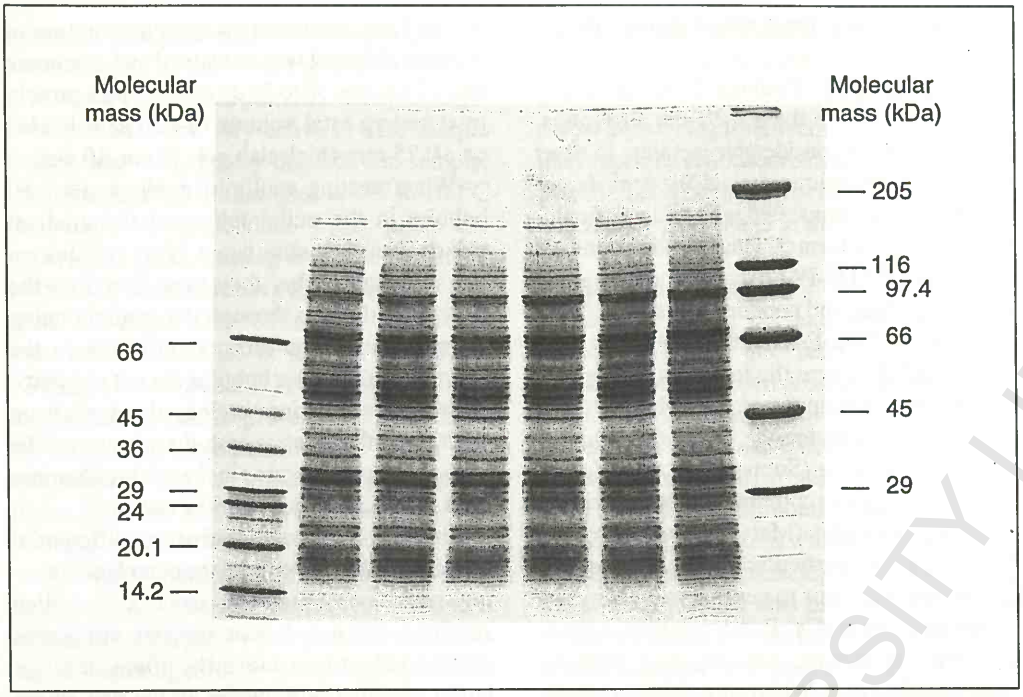


Figure 10.1.6 Separation of membrane proteins by 5.1% to 20.5% polyacrylamide gradient SDS-PAGE. Approximately 30 µl of 1× SDS sample buffer containing 30 µg of Alaskan pea (*Pisum sativum*) membrane proteins was loaded in wells of a 14 × 14–cm, 0.75-mm-thick gel. Standard proteins were included in the outside lanes. The gel was run at 4 mA for ~15 hr.

mensional SDS-PAGE should be considered (*UNIT 10.4*). Unusual protein compositions can cause anomalous mobilities during electrophoresis (see Critical Parameters and Troubleshooting), but similar-sized proteins of widely different amino acid composition or structure may still be resolved from one another using one-dimensional SDS-PAGE. Purified protein complexes or multimeric proteins consisting of subunits of different molecular size will be resolved into constituent polypeptides. Comparison of the protein bands obtained under nonreducing and reducing conditions provides information about the molecular size of disulfide cross-linked component subunits. The individual polypeptides can be isolated by electroelution (Smith, 1993) or electroblotting (*UNIT 10.8*), and the amino acid sequences can be determined.

Both the tricine (Schagger and von Jagow, 1987) and the modified Laemmli (Okajima et al., 1993) peptide separation procedures presented here (Alternate Protocols 1 and 2) are simple to set up and provide resolution down to 5 kDa. To separate peptides below 5 kDa, the tricine procedure must be modified by preparing a 16.5% T, 2.7% C resolving gel that uses a 10% T spacer gel between the stacking and resolving gel (Schagger and von Jagow, 1987). % C is the percentage of cross-linker (bisacrylamide) in the total monomer (acrylamide plus bisacrylamide).

Continuous electrophoresis, where the same buffer is used throughout the tank and gel, is popular because of its versatility and simplicity. The phosphate system described in Alternate Protocol 3 is based on that of Weber et al. (1972). Although unable to produce the high-resolution separations of the discontinuous SDS-PAGE procedures, continuous SDS-PAGE uses fewer solutions with one basic buffer and no stacking gel. Artifacts are also less likely to occur in continuous systems. Pepsin, for example, migrates anomalously on Laemmli-based discontinuous SDS-PAGE but has the expected mobility after electrophoresis in the phosphate-based continuous system described here. This is also true of cross-linked proteins.

Multiple gel casting (Support Protocols 1 to 3) is appropriate when gel-to-gel consistency is paramount or when the number of gels processed exceeds 5 a week. The variety of multiple gel casters, gradient makers, and inexpensive pumps available from major suppliers simplifies the process of casting gels in the laboratory. Alternatively, precast gradient gels

UNIVERSITY LIBRARY

are available for most major brands of gel apparatuses (Table 10.1.3).

Minigels (Basic Protocol 2) are generally considered to be in the 8 × 10–cm size range, although there is considerable variation in exact size. Every technique that is used on larger systems can be translated with little difficulty into the minigel format. This includes standard and gradient SDS-PAGE and separations for immunoblotting and peptide sequencing. Two-dimensional SDS-PAGE electrophoresis also adapts well, but here the limitation of separation area becomes apparent; for high-resolution separations, large-format gels are required. Gradient minigels (Support Protocol 3) are popular due to the combination of separation range and resolution (Matsudaira and Burgess, 1978). They are particularly useful for separation of proteins prior to peptide sequencing.

Mylar support (Gel Bond) provides a practical way of casting, running, and, staining extremely thin gels. When gels <0.75 mm thick are used, reagents have much better access both into and out of the gel, reducing staining time in both Coomassie blue and silver staining (*UNIT 10.5*). Double and broadened images caused by differential migration of the protein across the thickness of the gel are minimized, improving resolution.

Critical Parameters and Troubleshooting

If an electrophoretically separated protein will be electroeluted or electroblotted (*UNIT 10.8*) for sequence analysis, the highest-purity reagents available should be used. If necessary, SDS can be purified by recrystallization following the procedure given in Reagents and Solutions.

If the gels polymerize too fast, the amount of ammonium persulfate should be reduced by one-third to one-half. If the gels polymerize too slowly or fail to polymerize all the way to the top, fresh ammonium persulfate should be used or the amount of ammonium persulfate should be increased by one-third to one-half. The overlay should be added slowly down the spacer edge to prevent the overlay solution from crashing down and disturbing the gel interface.

After a separating gel is poured, it may be stored with an overlay of the same buffer used in the gel. Immediately prior to use, the stacking gel should be poured; otherwise, there will be a gradual diffusion-driven mixing of buffers between the two gels, which will cause a loss of resolution.

The protein of interest should be present in 0.2- to 1-μg amounts in a complex mixture of proteins if the gel will be stained by Coomassie blue. Typically, 30 to 50 μg of a complex protein mixture in a total volume of <20 μl is loaded on a 0.75-mm-thick slab gel (16 cm, 10 wells).

When casting multiple gradient gels, all bubbles in the outlet tubing of the gradient maker should be eliminated. If air bubbles get into the outlet tube, they may flow into the caster and then up through the gradient being poured, causing an area of distortion in the polymerized gel. Air bubbles are not so great a problem when casting single gradient gels from the top. As the gels are cast, the stirrer must be slowed so that the vortex in the mixing chamber does not allow air to enter the outlet.

Uneven heating of the gel causes differential migration of proteins, with the outer lanes moving more slowly than the center lanes (called smiling). Increased heat transfer eliminates smiling and can be achieved by filling the lower buffer chamber with buffer all the way to the level of the sample wells, by maintaining a constant temperature between 10° to 20°C, and by stirring the lower buffer with a magnetic stirrer. Alternatively, the heat load can be decreased by running at a lower current.

If the tracking dye band is diffuse, prepare fresh buffer and acrylamide monomer stocks. If the protein bands are diffuse, increase the current by 25% to 50% to complete the run more quickly and minimize band diffusion, use a higher percentage of acrylamide, or try a gradient gel. Lengthy separations using gradient gels generally produce good results (Fig. 10.1.6). Check for possible proteolytic degradation that may cause loss of high-molecular-weight bands and create a smeared banding pattern.

If there is vertical streaking of protein bands, decrease the amount of sample loaded on the gel, further purify the protein of interest to reduce the amount of contaminating protein applied to the gel, or reduce the current by 25%. Another cause of vertical streaking of protein bands is precipitation, which can sometimes be eliminated by centrifuging the sample or by reducing the percentage of acrylamide in the gel.

Proteins can migrate faster or slower than their actual molecular weight would indicate. Abnormal migration is usually associated with a high proportion of basic or charged amino acids (Takano et al., 1988). Other problems can occur during isolation and preparation of the protein sample for electrophoresis. Proteolysis of proteins during cell fractionation by endog-

FCU IAS CENTRAL UNIVERSITY LIBRARY

enous proteases can cause subtle band splitting and smearing in the resulting electrophoretogram (electrophoresis pattern). Many endogenous proteases are very active in SDS sample buffers and will rapidly degrade the sample; thus, first heating the samples to 70° to 100°C for 3 min is recommended.

In some cases, heating to 100°C in sample buffer will cause selective aggregation of proteins, creating a smeared layer of Coomassie blue–stained material at the top of the gel (Gallagher and Leonard, 1987). To avoid heating artifacts and also prevent proteolysis, the use of specific protease inhibitors during protein isolation and/or lower heating temperatures (70° to 80°C) have been effective (Dhugga et al., 1988).

Although continuous gels suffer from poor band sharpness, they are less prone to artifacts caused by aggregation and protein cross-linking. If streaking or aggregation appear to be a problem with the Laemmli system, then the same sample should be subjected to continuous SDS-PAGE to see if the problem is intrinsic to the Laemmli gel or the sample.

If the protein bands spread laterally from gel lanes, the time between applying the sample and running the gel should be reduced in order to decrease the diffusion of sample out of the wells. Alternatively, the acrylamide percentage should be increased in the stacking gels from 4% to 4.5% or 5% acrylamide, or the operating current should be increased by 25% to decrease diffusion in the stacking gel. Use caution when adding 1× SDS electrophoresis buffer to the upper buffer chamber. Samples can get swept into adjacent wells and onto the top of the well arm.

If the protein bands are uneven, the stacking gel may not have been adequately polymerized. This can be corrected by deaerating the stacking gel solution thoroughly or by increasing the ammonium persulfate and TEMED concentrations by one-third to one-half. Another cause of distorted bands is salt in the protein sample, which can be removed by dialysis (*APPENDIX 3B*), gel filtration (*UNIT 8.3*), or precipitation. Skewed protein bands can be caused by an uneven interface between the stacking and separating gels, which can be corrected by starting over and being careful not to disturb the separating gel while overlaying with isobutyl alcohol.

If a run takes too long, the buffers may be too concentrated or the operating current too low. If the run is too short, the buffers may be too dilute or the operating current too high.

If double bands are observed, the protein may be partially oxidized or partially degraded. Oxidation can be minimized by increasing the 2-ME concentration in the sample buffer or by preparing a fresh protein sample. If fewer bands than expected are observed and there is a heavy protein band at the dye front, increase the acrylamide percentage in the gel.

Anticipated Results

Polyacrylamide gel electrophoresis done under denaturing and reducing conditions should resolve any two proteins, except two of identical size. Resolution of proteins in the presence of SDS is a function of gel concentration and the size of the proteins being separated. Under nondenaturing conditions, the biological activity of a protein will be maintained.

Comparison of reducing and nonreducing denaturing gels can also provide valuable information about the number of disulfide cross-linked subunits in a protein complex. If the subunits are held together by disulfide linkages, the protein will separate in denaturing gels as a complex or as smaller sized subunits under nonreducing or reducing conditions, respectively. However, proteins separated on nonreducing denaturing gels appear more diffuse and exhibit less overall resolution than those separated on reducing gels.

Gradient gels provide superior protein-band sharpness and resolve a larger size range of proteins, making them ideal for most types of experiments in spite of being more difficult to prepare. Molecular-weight calculations are simplified because of the extended linear relationship between size and protein position in the gel. Increased band sharpness of both high- and low-molecular-weight proteins on the same gel greatly simplifies survey experiments, such as gene expression studies where the characteristics of the responsive protein are not known. Furthermore, the increased resolution dramatically improves autoradiographic analysis. Preparation of gradient gels is straightforward, although practice with gradient solutions containing dye is recommended. The gradient gels can be stored for several days at 0° to 4°C before casting the stacking gel.

Time Considerations

Preparation of separating and stacking gels requires 2 to 3 hr. Gradient gels generally take 5 min to cast singly. Casting multiple single-concentration gels requires an additional 10 min for assembly. Casting multiple gradient gels takes 15 to 20 min plus assembly time. It takes 4 to 5 hr to run a 14 × 14–cm, 0.75-mm

BOSTON UNIVERSITY LIBRARY

gel at 15 mA (70 to 150 V), and 3 to 4 hr to run a 0.75-mm gel at 20 mA (80 to 200 V). Overnight separations of ~12 hr require 4 mA per 0.75-mm gel. It takes 4 to 5 hr to run a 1.5-mm gel at 30 mA. Electrophoresis is normally performed at 15° to 20°C, with the temperature held constant using a circulating water bath. For air-cooled electrophoresis units, lower currents and thus longer run times are recommended.

It takes ~1 hr to run a 0.75-mm minigel at 20 mA (100 to 120 V). Separation times are not significantly different for gradient minigels.

Literature Cited

Dhugga, K.S., Waines, J.G., and Leonard, R.T. 1988. Correlated induction of nitrate uptake and membrane polypeptides in corn roots. *Plant Physiol.* 87:120-125.

Gallagher, S.R. and Leonard, R.T. 1987. Electrophoretic characterization of a detergent-treated plasma membrane fraction from corn roots. *Plant Physiol.* 83:265-271.

Hunkapiller, M.W., Lujan, E., Ostrander, F., and Hood, L.E. 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. *Methods Enzymol.* 91:227-236.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.

Matsudaira, P.T. and Burgess, D.R. 1978. SDS microslab linear gradient polyacrylamide gel electrophoresis. *Anal. Biochem.* 87:386-396.

Okajima, T., Tanabe, T., and Yasuda, T. 1993. Non-urea sodium dodecyl sulfate-polyacrylamide gel electrophoresis with high-molarity buffers for the separation of proteins and peptides. *Anal. Biochem.* 211:293-300.

Schagger, H. and von Jagow, G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166:368-379.

Sigma. Molecular weight markers for proteins kit (Technical Bulletin MWS-877L). Sigma Chemical Company, St. Louis, Mo.

Smith, J.A. 1993. Electroelution of proteins from stained gels. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 10.5.1-10.5.5. John Wiley & Sons, New York.

Takano, E., Maki, M., Mori, H., Hatanaka, N., Marti, T., Titani, K., Kannagi, R., Ooi, T., and Murachi, T. 1988. Pig heart calpastatin: Identification of repetitive domain structures and anomalous behavior in polyacrylamide gel electrophoresis. *Biochemistry* 27:1964-1972.

Weber, K., Pringle, J.R., and Osborn, M. 1972. Measurement of molecular weights by electrophoresis on SDS-acrylamide gel. *Methods Enzymol.* 26:3-27.

Key Reference

Hames, B.D. and Rickwood, D. (eds.) 1990. Gel Electrophoresis of Proteins: A Practical Approach, 2nd ed. Oxford University Press, New York.

An excellent book describing gel electrophoresis of proteins.

Contributed by Sean R. Gallagher
Hoefer Pharmacia Biotech
San Francisco, California

BCU IASI / CENTRAL UNIVERSITY LIBRARY

One-Dimensional Isoelectric Focusing of Proteins in Slab Gels

Separation of proteins in complex mixtures for further analysis or purification can be achieved by isoelectric focusing (IEF), a special type of electrophoresis in which movement of proteins through pores in a polyacrylamide gel matrix is modulated by the presence of a pH gradient created by soluble ampholytes. The types of reagents and equipment used are similar to those needed for standard Laemmli sodium dodecyl sulfate (SDS) gels (*UNIT 10.1*) and IEF tube gels (*UNIT 10.4*). Use of IEF in the slab-gel format provides the advantage of permitting the analysis of large numbers of samples in a single experiment.

The Basic Protocol describes isoelectric focusing of polypeptides in urea-containing polyacrylamide slab gels. A pH gradient forms when ampholytes (small charged organic molecules) migrate through the gel in response to an electric field. Denaturing of proteins in special solubilizing buffer allows maximum exposure of amino acid side chains that contribute to a protein's isoelectric point (pI), as well as dissociation of multisubunit proteins. Separation on the basis of pI values produces bands of polypeptides that can be electrotransferred to suitable membranes for further analysis, for example, by staining (*UNIT 10.8*) or immunodetection. If samples have been radiolabeled, they can be visualized by fluorography. This unit contains support protocols detailing special conditions for application of denaturing IEF slab gels for electrotransfer (Support Protocol 1) and fluorography (Support Protocol 2) in which a scintillator is incorporated into the gel after completion of the run to facilitate detection of radioactively labeled polypeptides.

ISOELECTRIC FOCUSING IN SLAB GELS UNDER DENATURING CONDITIONS

BASIC PROTOCOL

The preparation and running of an IEF slab gel containing nonionic detergent and urea is presented. An acrylamide solution supplemented with urea, Triton X-100, and ampholytes is used to cast a slab gel; the stacking gel normally used in protein electrophoresis is omitted. Protein samples (preferably salt-free) are solubilized in buffer containing urea, 2-mercaptoethanol (2-ME), a nonionic detergent such as Triton X-100, and ampholytes before electrophoresis. A variety of electrophoresis units may be used with this protocol, but gel lengths of 15 to 20 cm and thicknesses of 0.75 to 1.0 mm are recommended. Thicker gels can be run, but consume correspondingly larger quantities of expensive ampholytes, and cannot be run at high power at the onset of the run.

Soluble ampholytes (Hoefer Pharmacia, Oxford Glycosystems, Serva) in the desired pH range are commercially available. The choice of pH range is dictated by the isoelectric points of the proteins to be resolved. This protocol results in a rather broad pH gradient (~pH 4.5 to 7.5) and can be modified as needed, provided the final ampholyte concentrations (~2% w/v) are maintained as indicated.

Materials

Urea (any high-grade urea will usually prove satisfactory)
10% (v/v) Triton X-100
Acrylamide stock solution (see recipe)
Ampholytes pH 5-7
Ampholytes pH 3.5-10
10% (w/v) ammonium persulfate (store 500-µl aliquots of freshly prepared stock at –20°C)

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

TEMED (*N,N,N′-N′*-tetramethylenediamine)
20 mM H_3PO_4
Protein samples to be analyzed
IEF solubilization buffer (see recipe)
Overlay solution: IEF solubilization buffer diluted 1:3 (v/v) with H_2O
20 mM NaOH (freshly made, degassed immediately prior to use)

Electrophoresis apparatus (e.g., Studier type or equivalent) with glass plates, spacers, clamps, casting stand, and buffer chambers
Side-arm flask
Comb
Power supply (capable of delivering at least 1000 V and 10 W total power)
Hamilton syringe or mechanical pipet

CAUTION: Acrylamide is a neurotoxin; wear gloves when handling solutions.

Cast the isoelectric focusing gel

1. Prepare the electrophoresis apparatus by thoroughly cleaning glass plates. Assemble gel mold using spacers and clamps and mount in casting stand.

   Agarose (1% w/v) is a useful sealing agent. It is dispensed from a freshly boiled solution using a Pasteur pipet. Be sure that the hot agarose flows well between glass plates and spacer. Some minor spillage of agarose into the inside of the gel mold may be tolerated without affecting polymerization of the gel or quality of the separation.

   Most types of equipment for vertical slab gel electrophoresis can be used. Dimensions of gels are not specified, as these will differ widely between laboratories.

2. Prepare acrylamide gel solution in a side-arm flask; for 10 ml gel mix use the following:

   5.5 g urea (9.1 M final)
   2 ml 10% (v/v) Triton X-100 (2% v/v final)
   2 ml Milli-Q water
   1.35 ml acrylamide stock solution
   0.4 ml ampholytes pH 5-7
   0.1 ml ampholytes pH 3.5-10.

   To promote rapid dissolution of urea, the gel solution may be swirled in a 37°C water bath. Avoid overheating to minimize decomposition of urea.

   The total volume of the gel mix to be prepared is easily calculated from the dimensions of the gel and the thickness of the spacer. For a 200 × 150 × 1–mm gel, a total of 30 ml gel mix would be required; multiply quantities given above by three to obtain the appropriate amounts.

   The side-arm flask size (volume) should be at least 3× the volume of the amount of gel mix to be accommodated.

3. Degas solution briefly (up to 5 min).

   Because oxygen inhibits the polymerization process, it is important that the gel solution be deaerated before the polymerization catalysts are added.

4. Add 20 µl of 10% (w/v) ammonium persulfate and 10 µl TEMED to initiate polymerization. Mix well by swirling.

5. Pour gel all the way to the top of the glass plates.

   Wear gloves when handling unpolymerized gel solutions. Some of the gel mix will overflow from the mold. Remove spillage with tissue paper. By filling the mold to the

BOU IASI / CENTRAL UNIVERSITY LIBRARY

top, the insertion of the comb without causing bubbles to get trapped underneath the wells (step 6) is facilitated.

6. Insert comb, being careful to avoid air bubbles. Allow gel to polymerize ~45 min.

   The gel should be used shortly after polymerization is complete (the same day it is cast). In view of the decomposition of urea in aqueous solution, storage of poured gels is not recommended.

Perform the electrophoretic separation

7. Remove comb and place gel-plate sandwich in electrophoresis apparatus. Fill lower buffer chamber (anode) with 20 mM H_3PO_4.

8. Prepare protein samples to be analyzed in IEF solubilization buffer. Apply samples in gel wells. Fill any empty wells with a similar volume of IEF solubilization buffer.

   Samples may be applied with a Hamilton syringe or a mechanical pipet. The volume of the sample loaded is not particularly important, but it is recommended that no greater than 60% to 70% of the well's capacity be loaded so that overlay solution may be subsequently applied. Generally, sample volumes are <100 µl. Samples are loaded immediately adjacent to one another, without the need for empty lanes in between. Acrylamide teeth should be in good contact with both glass plates to eliminate the potential of spillover between wells. When in doubt, the integrity of the sample wells can be checked by first loading the overlay solution (which is colored); spillage from a freshly loaded well into an adjacent one indicates a faulty sample well, not to be used. If the wells are intact, samples can then be layered underneath the overlay using a Hamilton-type syringe.

   If unequal sample volumes are added to wells, the lane with the larger volume will spread during electrophoresis and constrict adjacent lanes, causing distortions. Preparing samples at approximately the same concentration and loading an equal volume in each well will ensure that all lanes are the same width and that the polypeptides run evenly. Fill any extra lanes (i.e., those not receiving a protein sample) with a similar volume of sample buffer, such that all wells are filled.

   The total amount of protein that may be loaded in a volume of 100 µl should not exceed 200 µg, and should be loaded over a total cross-section of 10 to 15 mm^2. The sample capacity of IEF gels is considerably greater than that of SDS-PAGE gels (UNIT 10.1), as it is an equilibrium separation method. The guideline given is recommended, but larger amounts of protein can in all likelihood be analyzed without serious loss of resolution. Samples should be salt-free, if possible, or of low ionic strength, to obtain reproducible separations. Immune complexes can be dissociated by the addition of IEF solubilization buffer. Most proteins in solution can be analyzed successfully when diluted with IEF solubilization buffer (4 parts buffer to 1 part sample). Samples that have been denatured in solubilization buffer for SDS-PAGE can likewise be diluted with IEF solubilization buffer and analyzed in parallel.

9. Apply overlay solution with a syringe or mechanical pipet, taking care not to disturb the sample layer.

   Use ~1/5 to 1/3 of the sample volume for the overlay. It is most convenient to add the tracking dye to the overlay solution, as this will generate a colored layer that separates sample from the upper reservoir electrolyte.

10. Using a Pasteur pipette, gently layer 20 mM NaOH over the samples, until the NaOH solution is flush with the top of the glass plate of the gel mold. Fill the upper reservoir with the remainder of the NaOH solution.

11. Connect power supply and apply voltage, starting at 20 V/cm. Limit voltage to 50 V/cm.

BOSTON UNIVERSITY LIBRARY

CAUTION: *The voltages and currents used during electrophoresis are dangerous and potentially lethal. Safety considerations are given in the introduction to UNIT 10.1.*

The cathode is at the top (NaOH-containing reservoir) and the anode is at the bottom (H_3PO_4-containing reservoir). The most convenient manner of running the gel is to use a power supply with a constant power setting and voltage limitation. The power is adjusted to a limit such that the desired starting voltage is obtained. Over the course of the run, the current will drop and the voltage will increase until the desired preset value is attained.

12. Run gel 13 to 16 hr (for a gel of a separating distance of ~20 cm).

    An overnight run is convenient. Over the course of the run, bromphenol blue included with samples will migrate the entire length of the gel and accumulate at the acidic end of the gel as a greenish-yellow band. When gels are run for longer times (>16 hr), no bromphenol blue may be discernible at all by the end of the run.

13. On termination of the run disconnect the power supply and disassemble the electrophoresis apparatus. Remove one of the glass plates from the gel-plate sandwich.

    The IEF gel will present a rippled appearance. The ripples correspond to positions of high local ampholyte concentrations that osmotically attract water. The ripple pattern is usually a good indicator of whether the gel ran straight and may allow one to gauge edge effects, "smiling" (curvature of migrating band), or the presence of samples with salt content drastically different from that of adjacent samples.

14. Process the gel, if desired, for staining (*UNIT 10.5*), electroblotting (see Support Protocol 1), or fluorography (see Support Protocol 2). Invert the glass plate with the gel still attached and hold it directly over the dish containing the appropriate solution. Using a spatula, peel off one corner of the gel, and continue loosening the gel until it begins to detach spontaneously from the glass plate. While holding the glass plate horizontally over the liquid-filled dish, minimize the distance that the gel will have to "fall" into the receptacle to reduce the risk of breakage.

SUPPORT PROTOCOL 1

ELECTROBLOTTING FROM DENATURING ISOELECTRIC-FOCUSING SLAB GELS

The Basic Protocol produces gels with separated polypeptides that can be subjected to standard electrotransfer procedures (*UNIT 10.7*). It is necessary, however, to remove the nonionic detergent (included in the IEF gel recipe) from the gel as it interferes with a number of commonly used staining protocols; the nonionic detergent also very significantly reduces the amount of protein transferred to membranes in electroblotting experiments. It is also necessary to bind SDS to the proteins. Both tasks are accomplished by repeated washings of the gel as follows: The gel is submerged in a shallow dish containing 50% methanol/1% (w/v) SDS/5 mM Tris·Cl (pH 8.0); the volume of wash solution should be 10× the volume of the gel. The gel is washed five times, 10 min each time, changing the solutions between washes. A convenient way to restrain the gel while decanting the used wash solution is to place paper towels or filter paper of a size that approximates that of the gel directly on the surface of the gel. Alternatively, liquid may be removed by aspiration; again, paper towels or filter paper may be used to hold the gel in the container. After the washes, the gel is ready for transfer as described in *UNIT 10.7*.

BCU IASRI CENTRAL UNIVERSITY LIBRARY

FLUOROGRAPHY OF DENATURING ISOELECTRIC-FOCUSING SLAB GELS

Slab gels containing separated polypeptides radiolabeled with ^{3}H, ^{14}C, or ^{35}S can be analyzed by fluorography. The use of the scintillant PPO (2,5-diphenyloxazole) is recommended to increase detectability of these weak β emitters and make the gels relatively easy to handle.

Materials

DMSO
DMSO/PPO solution (see recipe)
Polyacrylamide gel containing polypeptides of interest

Film cassette
Film
Intensifying screen

CAUTION: Solutions of PPO in DMSO readily permeate the skin, causing the PPO to precipitate subcutaneously. It is therefore imperative that gloves be worn at all times while handling DMSO/PPO solutions.

1. Fill a shallow dish with DMSO, using at least 10× the volume of the gel. Agitate for 20 min.
2. Remove DMSO by decantation or aspiration, and repeat step 1, using the same volume of DMSO.

   At this stage, the gel often presents an opaque appearance, with lines running perpendicular to the direction of the electrical field. We have not determined the identity of the material that forms this banding pattern, but it almost certainly relates to the distribution of the ampholytes. Like the "ripple" pattern alluded to in the Basic Protocol (step 13), it is a very good indicator of the quality of the separation and would reveal distortions due to overloaded lanes or high salt content in some of the samples.
3. Remove DMSO as in step 2, then add DMSO/PPO. Use at least 10× the volume of the gel. Agitate for 45 to 60 min.

   The DMSO/PPO solution may be reused multiple times and can be replenished by addition of the appropriate amount of PPO (assume that PPO consumed is equivalent to the gel volume). Spent DMSO/PPO solution may be diluted with water to precipitate remaining PPO. The PPO is then recovered by filtration and dried. In this manner, the (expensive) PPO can be used to maximum advantage.
4. Pour off and save DMSO/PPO. Add tap water (use at least 10× the gel volume, or as much as the container will allow) to precipitate the PPO in the gel, imparting mechanical strength to the gel. Agitate for 10 min.
5. Replace water. Wash at least two more times, replacing water at 20 min intervals.
6. Dry gel and place in film cassette. Cover with film, close cassette, and expose at –70°C.

   A brief (≤1 msec) exposure of the film to a flash of light from a photographic flash unit or stroboscope hypersensitizes the film, allowing a linear relationship to be drawn between the film image and the amount of radioactivity in the sample.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Acrylamide stock solution: 30% (w/v) acrylamide, 1.6% (w/v) bisacrylamide

Use the recipe in Table 10.1.1, increasing the amount of bisacrylamide to 1.6 g per 100 ml final volume. Any high-quality electrophoresis-grade acrylamide should be satisfactory. It may be advisable to treat the solution with activated charcoal. Use 5 g charcoal (Merck) per 500 ml acrylamide stock solution and remove the charcoal by filtration. Filter the final solution through a 0.45-μm filter. Store ≤30 days at 4°C in the dark.

DMSO/PPO solution

Dissolve 220 g PPO in 800 ml DMSO (technical-grade DMSO may be used). Store at room temperature.

IEF solubilization buffer

5.7 g urea
2.0 ml 10% (v/v) Triton X-100
0.2 ml ampholytes pH 3.5-10
0.5 ml 2-mercaptoethanol
H_2O to 10 ml final volume

Bromphenol blue may be added to samples at a final concentration of 0.01% (w/v) to improve visibility of proteins during loading and overlay. Alternatively, it may be added to the overlay solution.

COMMENTARY

Background Information

Electrophoretic methods for separating polypeptides are powerful analytical tools. Separation according to size is achieved by electrophoresis in polyacrylamide matrices in the presence of sodium dodecyl sulfate (SDS; see *UNIT 10.1*). Separation of polypeptides according to charge is accomplished by isoelectric focusing: a pH gradient stabilized by an electric field allows the polypeptides to migrate until they reach their isoelectric point. The combination of separation according to charge in one dimension and separation according to molecular weight in the second dimension has become a widespread analytical method capable of resolving complex protein mixtures (see *UNIT 10.4*).

Separation in a single dimension is informative when the complexity of the polypeptide mixture to be resolved is small, or when methods for detecting the specific polypeptide of interest are available. For example, when a complex mixture of proteins is resolved by separation under nondenaturing conditions (*UNIT 10.3*), gel overlay methods that measure enzyme activity or ligand-binding activity can be used to detect the protein of interest. The most commonly used separation methods, however, involve prior denaturation of the protein mixture, so that multisubunit proteins dissociate and each polypeptide migrates as a single species, according to its mass or charge. In this case, specific detection of polypeptides is usually accomplished by transferring the resolved polypeptides to a suitable membrane (*UNIT 10.7*) and using antibodies to detect the polypeptide of interest (*UNIT 10.10*), or lectins to reveal the position of glycoproteins.

In many instances the complexity of the mixture of proteins to be resolved is reduced by prior purification. In some cases, for example, radiolabeled protein samples are subjected to immunoprecipitation, and the proteins thus recovered are resolved electrophoretically. Other methods of labeling (e.g., biotinylation) could be used, bearing in mind that the attachment of label may affect the charge of the protein or result in the introduction of charge heterogeneity in an otherwise homogeneous sample.

When examining the post-translational modifications on a single polypeptide backbone (e.g., by glycosylation, Chapter 12, or phosphorylation, Chapter 13), it may be advantageous to analyze the samples by isoelectric focusing instead of SDS-PAGE, especially if

BCU IASLIBRARY CENTRAL UNIVERSITY LIBRARY

the modifications in question result in a charge difference (e.g., addition of sialic acids to protein-bound glycans; phosphorylation; and sulfation). Such charge changes are often difficult to detect by SDS-PAGE, because whereas SDS-PAGE is sensitive to charge differences in proteins to be resolved, such changes affect mobility in an often unpredictable manner, or result in very minor differences in migration. The comparison of multiple samples by two-dimensional methods, although feasible, is clearly a rather cumbersome procedure with the inherent pitfalls of variation between individual separations.

The protocol described here for one-dimensional IEF is an adaptation of the method developed by O'Farrell (1975) for two-dimensional gel electrophoresis. In O'Farrell's technique, the IEF gel is cast in a glass tube, and later serves as the starting sample for the second-dimension gel. We have adapted gels of a composition similar to those described by O'Farrell to the slab-format gels. Most equipment currently in use for SDS-PAGE should be satisfactory (*UNIT 10.1*). No cooling of the electrophoresis unit is required, as long as the power applied to the gel is held within reasonable limits. The technique as described here allows high resolution of membrane proteins and has been used to detect isoelectric variants of membrane proteins not separable by size, to detect N-linked glycan modifications such as the addition of sialic acids, and to detect phosphorylated species of membrane proteins (Eichholtz et al., 1992; Benaroch et al., 1995). The method also lends itself to immunoblotting procedures, provided the high Triton X-100 concentrations in the gel are reduced and SDS is added by extensive washing of the gel in a methanol/SDS-containing solution. The method described is particularly useful when comparing polypeptides that may show only minute differences in charge: such differences would be more difficult to detect on two-dimensional separations, unless sophisticated methods are used to standardize gel pouring, electrophoresis, and workup of the gels (see Garrels, 1989). In addition, this method may be used to resolve short peptides, provided they can be detected (radiolabeled peptides).

Critical Parameters and Troubleshooting

It is important that the samples to be analyzed be salt-free, if possible. For immune complexes, this is easily accomplished by a final rinse of the complex adsorbed onto protein A–Sepharose, or some similar matrix, with water. Dissociation of the samples from the Sepharose is then achieved by addition of IEF solubilization buffer.

Staining of IEF gels as described in this protocol requires that the proteins are initially fixed in the gel using 20% (w/v) TCA (note that exposure of skin to TCA can cause blistering and other damage). This is followed by repeated washes in 10% (w/v) acetic acid/30% methanol. The gel can then be stained with most conventional SDS-PAGE stains (see *UNIT 10.5*).

Given the extreme variability of available electrophoresis equipment, one consideration is the size of the gel to be used. We have found a length of 15 to 20 cm to be satisfactory. Shorter gels may induce loss of the gel: the gel may simply slide from between the glass plates, as the combination of Triton and urea in the gel makes for poor adhesion to the glass plates. To avoid this, one may increase the surface area of the gel (pour larger gels), or assemble the gel mold with spacers such that the gel produced will taper towards the bottom (i.e., 1 cm narrower at the bottom than at the top).

Occasionally one may find that a gel has polymerized in an irregular fashion. Because the acrylamide gel serves as an anticonvection medium to stabilize the pH gradient, and does not contribute to resolution of the polypeptide mixture, such irregularities are usually inconsequential.

If the gel does not polymerize all the way to the top and hence no depressions of adequate size are present to form the sample wells, one may prepare a small batch of fresh gel solution and repour the top part separately. In this case, care should be taken to remove any remaining liquid from the top of the first gel, using blotting paper. The concentrations of ammonium persulfate and TEMED may be increased twofold to obtain rapid polymerization of the second, top layer. Any discontinuity in the gel that results from this procedure is unlikely to affect the separation. In this manner, an initially defective gel can be rescued with minimal loss of costly ampholytes.

Samples in SDS-PAGE solubilization buffer may be used for IEF, provided the same amount of buffer is present in each lane of the gel. Try to keep the amount of SDS sample buffer to be loaded as low as possible (30 µl or less). Dilute these samples with IEF solubilization buffer as much as the capacity of the sample well will allow (preferably more than a twofold dilution), and make sure that every sample well in the gel will receive the same

SCANNED BY RAE INTERNATIONAL LIBRARY

amount of SDS sample buffer (including lanes that will not receive protein samples).

Polypeptides that are rather basic will be poorly resolved on the gradients specified in this protocol. This may be compensated, but only in part by using ampholytes of a higher pH range. Although so-called nonequilibrium IEF gels, in which the samples are loaded at the acidic end of the gel and the run is terminated before equilibrium has been reached, have been successful when run in glass tubes, we have generally not been successful in adapting slab gels to nonequilibrium IEF conditions, except for runs of very short duration to resolve peptides (Heemels et al., 1993).

When using equipment that can be cooled, care should be taken not to cool the gels to the point where urea comes out of solution in the gel. Nonetheless, it is recommended that the temperature of the gel not be allowed to rise over 40°C (i.e., warm to the touch). Even though the high concentration of ampholytes in the gel is likely to scavenge most of the decomposition products, decomposition of urea in aqueous solution accelerates at higher temperatures and is a cause for concern (i.e., results in carbamylation of proteins).

For fluorography of gels, the use of the DMSO/PPO protocol is recommended, because the final product, the PPO-impregnated gel, is easier to handle and less fragile than gels processed using commercially available fluorography reagents such as EN^{3}HANCE (DuPont NEN) and similar products.

Anticipated Results

For representative examples of the types of separation that can be achieved using these protocols, the reader is referred to Eichholtz et al. (1992) and Benaroch et al. (1995).

Time Consideration

Preparing the gel mix, pouring the gel, and completing the polymerization process can be accomplished in <2 hr. Gels ~20 cm in length are conveniently electrophoresed overnight. Because the separation is essentially taken to equilibrium, variations in running time have minor impact on the resolution actually achieved. Runs should not be extended beyond 18 hr, however, as cathodic drift may begin to seriously affect the pH gradient in the basic portion of the gel. The time required for electroblotting will depend somewhat on the conditions used, but preparation of the gel for transfer should take 1 hr. The entire fluorography procedure requires 2.5 to 3 hr, not including drying of the gel, which can take 0.5 to 1 hr depending on the equipment used.

Literature Cited

Benaroch, P., Yilla, M., Raposo, G., Ito, K., Miwa, K., Geuze, H.J., and Ploegh, H.L. 1995. How MHC Class II molecules reach the endocytic pathway. *EMBO J.* 14:37-49.

Eichholtz, T.E., Vossebeld, P., Van Overveld, M., and Ploegh, H.L. 1992. Activation of protein kinase C accelerates internalization of transferrin receptor but not major histocompatibility complex class I molecules, independent of their phosphorylation status. *J. Biol. Chem.* 267:22490-22495.

Garrels, J.I. 1989. The QUEST system for quantitative analysis of two-dimensional gels. *J. Biol. Chem.* 264:5269-5282.

Heemels, M.T., Schumacher, T.N.M., Wonigeit, K., and Ploegh, H.L. 1993. Different peptides are translocated by allelic variants of the transporter associated with antigen presentation (TAP). *Science* 262:2059-2063.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.

Key References

Benaroch, et al., 1995. See above.

Eichholtz et al., 1992. See above.

These contain representative examples of good quality separations of membrane proteins attained by the IEF methods described here.

Contributed by Hidde L. Ploegh
Massachussetts Institute of Technology
Cambridge, Massachusetts

BOU IASTY CENTRAL UNIVERSITY LIBRARY

One-Dimensional Electrophoresis Using Nondenaturing Conditions

Nondenaturing or "native" electrophoresis—i.e., electrophoresis in the absence of denaturants such as detergents and urea—is an often-overlooked technique for determining the native size, subunit structure, and optimal separation of a protein. Because mobility depends on the size, shape, and intrinsic charge of the protein, nondenaturing electrophoresis provides a set of separation parameters distinctly different from mainly size-dependent denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE; *UNITS 10.1 & 10.4*) and charge-dependent isoelectric focusing (IEF; *UNITS 10.2 & 10.4*). Two protocols are presented below. Continuous PAGE (Basic Protocol) is highly flexible, permitting cationic and anionic electrophoresis over a full range of pH. The discontinuous procedure (Alternate Protocol) is limited to proteins negatively charged at neutral pH but provides high resolution for accurate size calibration.

CONTINUOUS ELECTROPHORESIS IN NONDENATURING POLYACRYLAMIDE GELS

BASIC PROTOCOL

Separation of proteins by nondenaturing electrophoresis requires the same type of equipment used for denaturing slab gels (*UNIT 10.1*) and is adaptable to a range of gel sizes (e.g., from 7.3 × 8.3–cm minigels to 14 × 16–cm full-size gels) and matrix types (e.g., single-concentration and gradient gels). This protocol outlines straightforward procedures for making acrylamide solutions, casting separating gels (stacking gels are omitted), loading samples, and conducting electrophoresis. Continuous systems, although flexible, do not give the high-resolution separation found in discontinuous systems (see Alternate Protocol).

Separation in a continuous system (i.e., in which the same buffer is used for preparing acrylamide solutions and filling electrophoresis chambers) is governed by pH, and this protocol describes four types of buffers useful over discrete ranges from pH 3.7 to pH 10.6. Use of unadjusted acetic acid gel buffer can extend the range to pH 2.0. The choice of pH and thus the buffer system depends on the protein being studied (i.e., its isoelectric point) and often must be determined empirically. In general, the system should be between pH 5.0 and 8.0 for optimal results. Extremes of pH can lead to precipitation or denaturation of the protein. Acrylamide concentrations are empirically determined, but the higher the percent acrylamide, the sharper the protein bands.

It is important to include native protein standards in the electrophoresis runs. Several manufacturers supply standards for isoelectric focusing that are also suitable for native electrophoresis. The standards have a range of isoelectric points and will carry a net positive, negative, or zero charge depending on the pH of the gel system. Alternatively, Sigma supplies a standard kit that is useful for calculating molecular weights under neutral pH, nondenaturing conditions. The samples and standard proteins, should be used at concentrations of ~1 to 2 μg/μl.

Materials

4× acetic acid gel buffer (200 mM acetic acid, pH 3.7 to 5.6; see recipe)
4× phosphate gel buffer (400 mM sodium phosphate, pH 5.8 to 8.0; see recipe)
4× Tris gel buffer (200 mM Tris·Cl, pH 7.1 to 8.9; see recipe)
4× glycine gel buffer (200 mM glycine, pH 8.6 to 10.6; see recipe)
300 mM sodium sulfite (0.38 g in 10 ml H_2O; used in acetic acid gel preparation)
Protein samples to be analyzed
Native protein standards

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON COLLEGE UNIVERSITY LIBRARY

Electrophoresis buffer: appropriate 4× gel buffer diluted to 1× with H_2O

75-ml side-arm flask (used in gel preparation)

Additional reagents and equipment for gel electrophoresis (*UNIT 10.1*) and staining proteins in gels (*UNIT 10.5*)

Prepare the gel

1. Assemble the glass-plate sandwich of the gel electrophoresis unit and secure it to the casting stand.

   Either single-concentration or gradient gels can be used in the minigel or standard-size format. Gradient gels will enhance the band sharpness of the separated proteins.

2. Prepare acrylamide solutions according to the recipes in Table 10.3.1, Table 10.3.2, Table 10.3.3, or Table 10.3.4, adding the ammonium persulfate and TEMED just before use.

   Deaeration of the solution before the polymerization catalysts are added will speed polymerization by removing inhibitory oxygen, but is not generally required. The pH used depends on many factors. The most important are the pI values of both the protein of interest and any contaminants, as well as protein mobility and protein solubility. Determining which pH and thus which buffer system to use is largely empirical. However, extremes of pH (<4.0 and >9.0) can lead to denaturation and should be avoided. Prior knowledge of the pI of a protein (UNITS 10.2 & 10.4) allows determination of the net charge under the separation conditions (i.e., if gel pH < protein pI, the protein

Table 10.3.1 Recipes for Acetic Acid Nondenaturing Polyacrylamide Gels[a]: pH range 3.7 to 5.6[b]

Stock solution[c]	Final acrylamide concentration in gel (%)[d]						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
300 mM sodium sulfite[e]	0.4	0.4	0.4	0.4	0.4	0.4	0.4
4× acetic acid gel buffer	10	10	10	10	10	10	10
H_2O	22.58	19.28	15.98	12.48	9.28	5.96	2.68
10% (w/v) ammonium persulfate[e]	0.3	0.3	0.3	0.3	0.3	0.3	0.3
TEMED[f]	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 300 mM sodium sulfite, 4× acetic acid gel buffer (see Reagents and Solutions), and H_2O. If desired to speed polymerization, degas under vacuum ~5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

[a]The recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

[b]The pH range can be extended to ~2.0 (the pH of acetic acid) by using unadjusted acetic acid in place of 4× acetic acid gel buffer, although there is little buffering capacity at this pH.

[c]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[d]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

[e]Must be freshly made. Sodium sulfite is needed for efficient polymerization at acid pH.

[f]Added just before polymerization.

RGU IAST / CENTRAL UNIVERSITY LIBRARY

Table 10.3.2 Recipes for Phosphate Nondenaturing Polyacrylamide Gels[a]: pH range 5.8 to 8.0

Stock solution[b]	Final acrylamide concentration in gel (%)[c]						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× phosphate gel buffer	10	10	10	10	10	10	10
H_2O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate[d]	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED[e]	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 4× phosphate gel buffer (see Reagents and Solutions), and H_2O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

[a]The recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

[d]Must be freshly made.

[e]Added just before polymerization.

will have a net positive charge; if gel pH > protein pI, the protein will be negatively charged).

3. Pour gel to 2 cm from the top of the gel mold and insert the comb. Avoid trapping air bubbles under the comb teeth.

   Air bubbles will cause small semicircular depressions in the well and lead to distortions in the protein banding pattern.

4. Allow gel solution to polymerize 1 to 2 hr.

   Polymerization is indicated by a sharp optical discontinuity around the wells.

Prepare samples and load the wells

5. Solubilize the protein sample to be analyzed using 5% (w/v) sucrose in water or dilute (1 to 5 mM) gel buffer if possible. Also prepare native protein standards.

   The concentration of protein will vary depending on the sample complexity and detection method. For Coomassie blue staining of highly enriched samples such as the standards, use 1 to 2 mg/ml (1 to 2 μg/μl). For more complex mixtures, use 5 to 10 mg/ml (5 to 10 μg/μl). Load 10- to 100-fold less for silver staining. In general, samples should be loaded in a minimum volume, preferably 10 to 20 μl for 0.75- and 1.5-mm-thick gels, respectively. With thin gels, this means using a more concentrated protein sample.

6. Remove comb carefully and rinse wells with electrophoresis buffer (appropriate 4× gel buffer diluted to 1×).

   Rinsing with electrophoresis buffer is needed to remove residual unpolymerized acrylamide monomer, which will continue to polymerize after comb removal, creating uneven wells that may interfere with sample loading.

BOSTON UNIVERSITY MEDICAL CENTRAL LIBRARY

Table 10.3.3 Recipes for Tris Nondenaturing Polyacrylamide Gels[a]: pH range 7.1 to 8.9

Stock solution[b]	Final acrylamide concentration in gel (%)[c]						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× Tris gel buffer	10	10	10	10	10	10	10
H_2O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate[d]	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED[e]	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 4× Tris gel buffer (see Reagents and Solutions), and H_2O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

[a]The recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

[d]Must be freshly made.

[e]Added just before polymerization.

7. Fill wells with electrophoresis buffer. If desired, prerun gel.

The gel can be prerun at this point to remove any charged material such as ammonium persulfate from the gel prior to loading the sample. Assemble the electrophoresis unit and fill the buffer chambers with electrophoresis buffer. Run the gel at 300 V until the current no longer drops. This should take ~30 min. Disassemble the unit, discard the buffer, and proceed to the next step.

8. Carefully load up to 10 µl (0.75-mm gels) or 20 µl (1.5-mm gels) sample per lane as a thin layer at the bottom of the wells. Load control wells with native protein standards. Add an equal volume of electrophoresis buffer to any empty wells to prevent spreading of adjoining lanes.

Mobility (R_f) markers require special consideration in nondenaturing gel systems. For cationic systems, cytochrome c (pI ~9 to 10, 5 to 10 µg/lane) works well as an R_f marker. Bromphenol blue (10 µg/ml) is a suitable marker for anionic systems. The marker should be included in the solubilization buffer with the sample.

Perform the separation

9. Assemble the gel unit, fill the upper and lower buffer chambers with electrophoresis buffer, and connect the unit to the power supply. Set current to 30 mA for a 1.5-mm-thick gel (15 mA for a 0.75-mm-thick gel).

If the protein is negatively charged under the separation conditions, then the standard SDS-PAGE electrode polarity should be used (proteins will migrate to the anode or positive electrode). If the protein is positively charged, then the electrodes should be reversed at the power supply (i.e., red high-voltage cable to the black output and black high-voltage lead to the red output) so the positively charged protein migrates to the negative cathode.

RGU IASI / CENTRAL UNIVERSITY LIBRARY

Table 10.3.4 Recipes for Glycine Nondenaturing Polyacrylamide Gels[a]: pH range 8.6 to 10.6

Stock solution[b]	Final acrylamide concentration in gel (%)[c]						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× glycine gel buffer	10	10	10	10	10	10	10
H_2O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate[d]	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED[e]	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.1.1), 4× glycine gel buffer (see Reagents and Solutions), and H_2O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

[a]The recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

[b]All reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

[c]Units of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

[d]Must be freshly made.

[e]Added just before polymerization.

10. Continue electrophoresis until the R_f marker reaches the bottom of the gel.

For minigels, electrophoresis will take 1 to 2 hr. Standard gels require 4 to 6 hr runs.

11. Turn off power supply, disassemble the unit, and remove gel from sandwich.

12. Stain the gel according to *UNIT 10.5*.

NATIVE DISCONTINUOUS ELECTROPHORESIS AND GENERATION OF MOLECULAR WEIGHT STANDARD CURVES (FERGUSON PLOTS)

ALTERNATE PROTOCOL

One straightforward approach to discontinuous native electrophoresis is to leave out the SDS and reducing agent (DTT) from the standard Laemmli SDS-PAGE protocol (*UNIT 10.1*). The gels are prepared as described in *UNIT 10.1* except that the sample buffer contains no SDS or DTT (samples are not heated), and the gel and electrophoresis solutions are prepared without SDS. This protocol illustrates the separation of standard proteins at four different concentrations of acrylamide and how the results are used to construct a molecular weight standard curve (Ferguson plot) without the need for SDS. By plotting relative mobility against %T (percentage weight per volume of acrylamide plus bisacrylamide in the gel), the presence of isoforms and multimeric proteins can also be detected.

Materials

4× Tris·Cl, pH 8.8 (1.5 M Tris·Cl; *APPENDIX 2E*)
4× Tris·Cl, pH 6.8 (0.5 M Tris·Cl; *APPENDIX 2E*)
Protein sample of interest
2× Tris/glycerol sample buffer (see recipe)

RCU IASH / CENTRAL UNIVERSITY LIBRARY

Native protein standards (e.g., Sigma nondenatured protein molecular weight kit)
Tris/glycine electrophoresis buffer (see recipe)

1. Assemble the glass-plate sandwich of the gel electrophoresis unit and place it in the casting stand.
2. Prepare and cast the gels, using 4× Tris·Cl, pH 8.8, for the separating gel and 4× Tris·Cl, pH 6.8, for the stacking gel instead of the SDS-containing counterparts (Table 10.1.1). Prepare a minimum of four separate gels at different acrylamide concentrations.

   A typical range of concentrations is from 5% to 12.5% (e.g., 5%, 7.5%, 10%, 12.5% acrylamide). As with SDS-PAGE, typical gel thickness ranges from 0.75 to 1.5 mm. The 0.75-mm-thick gels are recommended because they offer a combination of fast staining and high resolution.

3. Mix protein sample of interest 1:1 with 2× Tris/glycerol sample buffer to attain a 1 to 2 µg/µl final concentration. Also prepare native protein standards. Remove comb, rinse wells, and load 10 to 20 µl per well for Coomassie brilliant blue staining, and 1 to 2 µl for silver staining.

   Some proteins must be dissolved in 50 mM NaCl or water to become fully solubilized prior to mixing with the sample buffer (Sigma, 1986).

4. Assemble gel electrophoresis unit, using Tris/glycine electrophoresis buffer to fill both lower and upper buffer chambers. Connect power supply and conduct electrophoresis.

   Conditions for separation are the same as for discontinuous SDS-PAGE (i.e., 30 mA for 1.5-mm-thick gels, 15 mA for 0.75-mm-thick gels). For standard-size gels the separation takes 4 to 5 hr; for minigels, 1 to 2 hr is required. Alternatively, standard gels can be run at 4 to 6 mA/gel overnight.

5. After the bromphenol blue R_f marker has reached the bottom of the gel, fix and stain the proteins in the gels according to *UNIT 10.5*. Estimate relative mobilities of the proteins.

   An example of a stained gel is shown in Figure 10.3.1. A minimum of four gel concentrations is recommended. In Figure 10.3.1, Sigma native molecular weight standards were separated on 5%, 7.5%, 10%, and 12.5% acrylamide gels (5.1%, 7.7%, 10.3%, and 12.8% T, respectively).

6. Plot log R_f against gel concentration (% T) (Fig. 10.3.2). Determine the slope of K_r using linear regression.
7. Plot $-\log K_r$ of the curves from step 6 against log molecular weight of the standards (Fig. 10.3.3). Determine the slope using linear regression.
8. Estimate the size of the standards and unknowns from the generated curve (Ferguson plot).

   Use the curve generated by linear regression to estimate the predicted size of the standard for comparison to the actual size stated by the supplier. This indicates the accuracy of the curve. The $-\log K_r$ value (y) of the unknown is then used to predict the molecular weight (x).

RGU IASI/CENTRAL UNIVERSITY LIBRARY

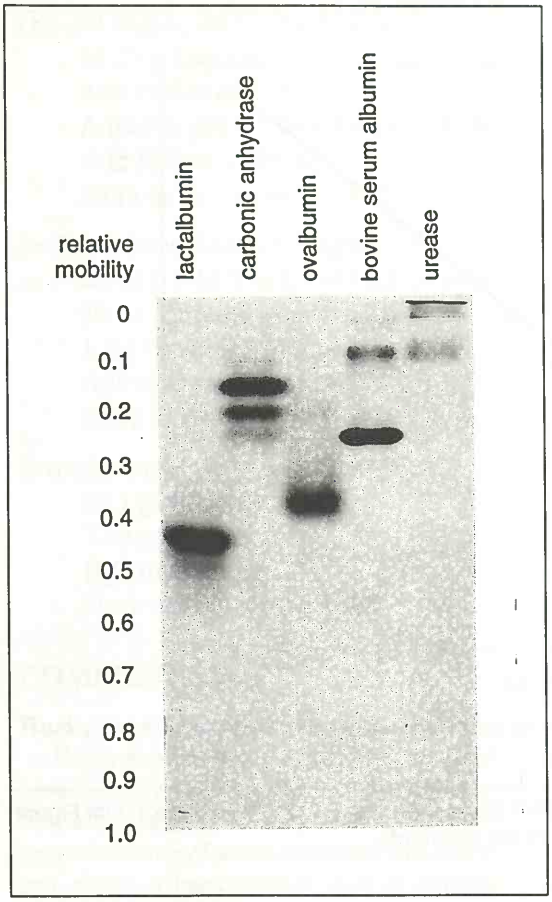


Figure 10.3.1 Separation of native protein standards under nondenaturing conditions by discontinuous polyacrylamide gel electrophoresis at 12.8% T. Approximately 20 µg protein was loaded per lane on a 1.5-mm-thick, 16-cm-long gel. The gel and samples were prepared according to the Alternate Protocol and were electrophoresed 16 hr at 6 mA. Proteins were stained with Coomassie blue.

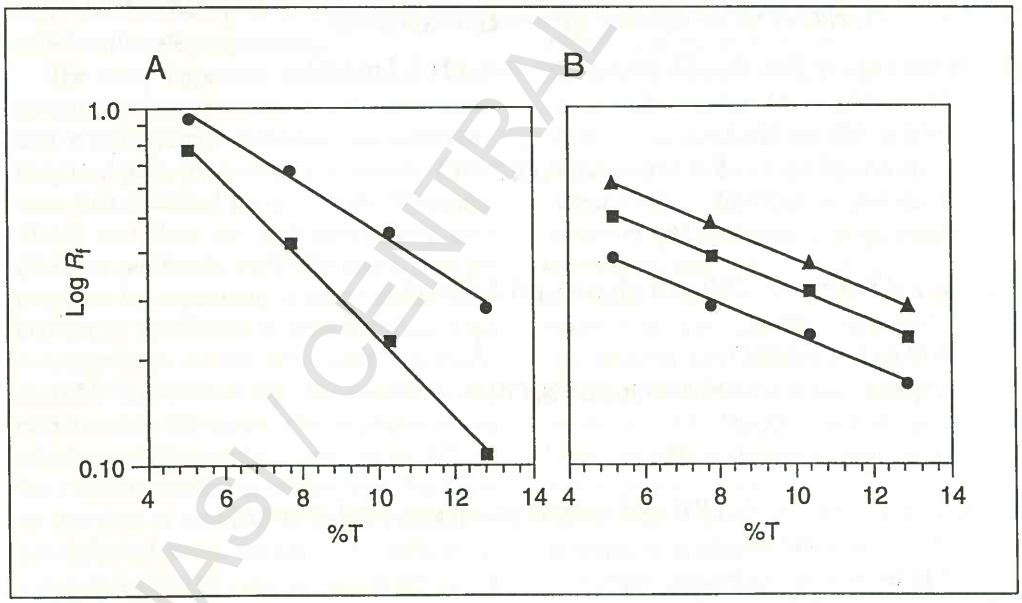


Figure 10.3.2 Effect of %T on the relative mobility of several native proteins. The relative mobility (R_f) of the standard proteins shown in Figure 10.3.1 was determined at four different gel concentrations and plotted as log R_f against %T. See text for details. (**A**) BSA monomer (squares) and dimer (circles); (**B**) carbonic anhydrase isoforms.

BOKHARI / CENTRAL UNIVERSITY LIBRARY

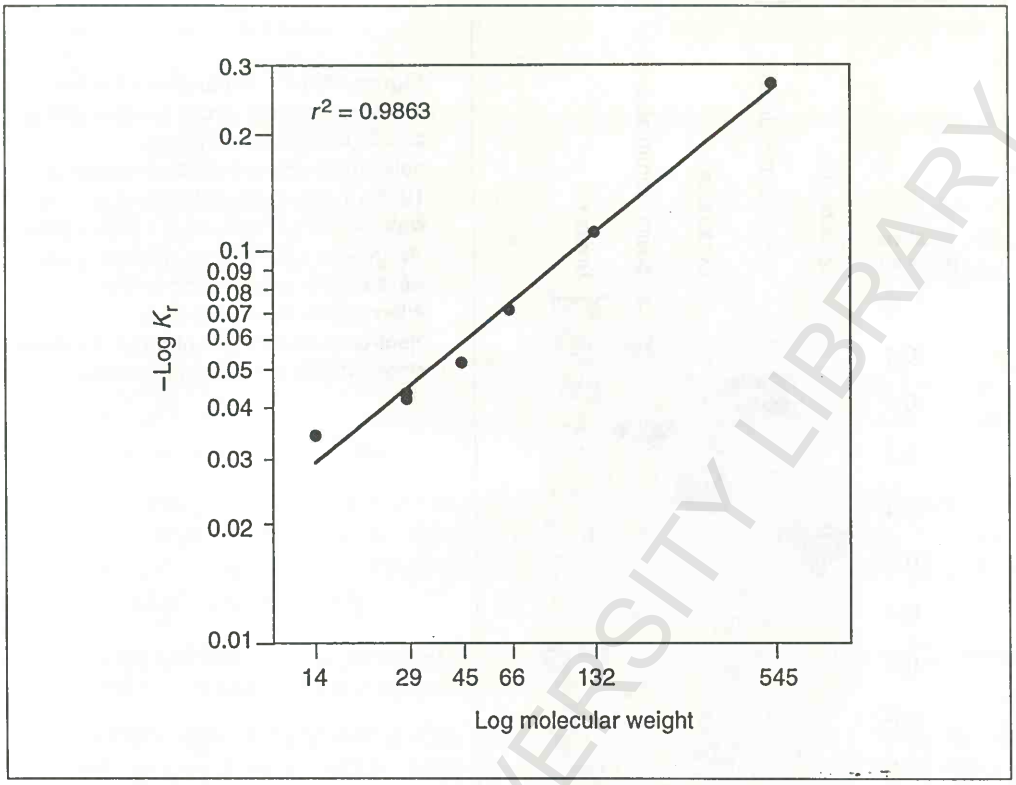


Figure 10.3.3 Native molecular weight standard curve. The –log slope of the line (K_r) from Figure 10.3.2 is plotted against log molecular weight of the standards.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Acetic acid gel buffer, 4× (200 mM acetic acid, pH 3.7 to 5.6)

11.49 ml glacial acetic acid
Add to 500 ml H_2O
Adjust to pH 3.7 to 5.6 with 1 M NaOH
Add H_2O to 1000 ml
Store up to 1 month at 4°C

Glycine gel buffer, 4× (200 mM glycine, pH 8.6 to 10.6)

15.01 g glycine
Add to 500 ml H_2O
Adjust to pH 8.6 to 10.6 with 1 M NaOH
Add H_2O to 1000 ml
Store up to 1 month at 4°C

Phosphate gel buffer, 4× (400 mM sodium phosphate, pH 5.8 to 8.0)

55.2 g $NaH_2PO_4 \cdot H_2O$
Add to 500 ml H_2O
Adjust to pH 5.8 to 8.0 with 1 M NaOH
Add H_2O to 1000 ml
Store up to 1 month at 4°C

BOU IASL / JEN E LUNIVERSITY LIBRARY

Tris gel buffer, 4× (200 mM Tris·Cl, pH 7.1 to 8.9)

24.23 g Tris base
Add to 500 ml H_2O
Adjust to pH 7.1 to 8.9 with 1 M HCl
Add H_2O to 1000 ml
Store up to 1 month at 4°C

2× Tris/glycerol sample buffer

25 ml 0.5 M Tris·Cl, pH 6.8 (*APPENDIX 2E*)
20 ml glycerol
1 mg bromphenol blue
Add H_2O to 100 ml and mix
Store in 1-ml aliquots up to 6 months at –70°C

Tris/glycine electrophoresis buffer

15.1 g Tris base
72.0 g glycine
H_2O to 5000 ml
Store up to 1 month at 4°C

COMMENTARY

Background Information

Under nondenaturing conditions, in which protein activity, native charge, and conformation are sustained, electrophoretic separation depends on many factors, including size, shape, and charge. Characteristics such as intrinsic molecular weight (i.e., in the absence of denaturation), the number of isoforms, and the presence of multimeric proteins can be determined with nondenaturing electrophoresis (often called native electrophoresis).

The most important application of nondenaturing electrophoresis is in the determination of native protein size (Alternate Protocol). Ferguson plots (reviewed by Andrews, 1986) were first described for starch gels (Ferguson, 1964) and then for polyacrylamide gels (Hedrick and Smith, 1968). Ferguson plots are prepared by separating proteins under nondenaturing conditions at several different gel concentrations. As the acrylamide concentration (%T) is increased, the relative mobility (R_f) of the protein decreases. This is plotted as log relative mobility (on the *y* axis) versus %T (on the *x* axis) to produce a straight line. The slope of this line is referred to as the retardation coefficient (K_r) and measures how effectively a protein is slowed by the increase in %T. Large proteins will be retarded much more significantly than small proteins with increasing gel concentration, with the size of the protein being proportional to the slope of the curve. Once the K_r plots for several size standards are generated (Fig. 10.3.2), the K_r values are plotted against the molecular weight of the standard proteins using a log-log graph (Fig. 10.3.3). The retardation coefficient also depends on a large number of other variables including temperature, pH, buffer type, ionic strength, and %C (percent bisacrylamide cross-linker). All these factors should be kept constant for a given experiment.

In addition to estimated size, other types of information are available from the Ferguson plots (Rodbard and Chrambach, 1971; Andrews, 1986). For example, if two components differ in size but have the same charge per unit size (e.g., for a multimeric protein with identical subunits), curves similar to those illustrated by BSA monomer and dimer (Fig. 10.3.2A) will result. Note that when the curve is extrapolated back to 0% T, it is evident that the monomer and the dimer have similar free solution mobilities. Furthermore, as the acrylamide concentration is increased, the separation between the two also increases. However, if two proteins have similar sizes but different amounts of charge, the curves will be parallel on the log plot. This is illustrated by the carbonic anhydrase isoforms (Fig. 10.3.2B). In this example, optimal separation of the isoforms occurs at the lower concentrations of acrylamide as this is a log plot.

Further applications of nondenaturing electrophoresis include preparative purification. The pH of the gel determines the net charge on the protein. Below its isoelectric point (pI) a protein will have a net positive charge, whereas above its pI it will have a net negative charge. In general, most proteins will be positively charged at pH 2.0 to 4.0; above pH 8.0, most

B. C. U. „M. EMINESCU" IAŞI
CHILIE

proteins will be negatively charged. As these general guidelines imply, the majority of proteins have isoelectric points between pH 4.0 and 8.0. There are, however, many exceptions. A protein with a highly acidic isoelectric point (e.g., pepsin, with a pI of 2.2) will remain negatively charged at a pH down to its pI. Although a full range of pH options are given, extremes of pH (<4.0 and >9.0) should be avoided, if possible, to minimize denaturation or inactivation. By picking an appropriate electrophoresis pH, it is possible to ensure that the protein of interest will be either positively or negatively charged so that it can be selectively run into the gel, excluding a large proportion of contaminants that have the opposite or no charge. Furthermore, the pH conditions determine the mobility and can be adjusted to ensure a difference in mobility between the protein of interest and contaminants.

Continuous gel systems (Basic Protocol) offer the most flexibility in terms of separation design. The pH can be tailored so that a given protein has a net positive, neutral, or negative charge. Depending on the polarity of the gel, the protein can then be excluded from or electrophoresed into the gel. Discontinuous gels have a fixed pH and gel polarity. For the nondenaturing Laemmli gel presented in the Alternate Protocol, the proteins of interest should have an isoelectric point of ≤7.0 in order to be negatively charged so that they move into the gel. Other more basic and more acidic discontinuous gel systems can be found in Hames (1990) and Schägger (1994).

Critical Parameters

The success of a gel separation under nondenaturing conditions depends on many factors, and two of the most important are protein solubility and isoelectric point. The protein must be soluble at the pH and the ionic strength of the gel, and it must be charged at that pH in order to move into the gel. If the protein experiences a pH below its isoelectric point, then it will have a net positive charge and will move to the negative electrode. Note that this is the reverse of typical SDS-PAGE. If the protein experiences a pH above its isoelectric point, it will have a net negative charge and will migrate to the positive electrode.

Solubility is a complex issue. Membrane-associated and other hydrophobic proteins are difficult to separate by nondenaturing electrophoresis (Schägger, 1994). Nonionic detergents at concentrations up to 1% and solubilizing reagents such as urea (4 to 8 M) can be used, but these reagents, especially urea, are likely to alter the protein's conformation and most likely the isoelectric point by exposing previously hidden charged groups. If detergent or urea must be included for solubilization, the minimum required to solubilize the protein should be used. Schägger (1994) lists several nonionic detergents suitable for solubilization. Among the more popular are octylglucoside and CHAPS. In general, detergents should be used near the critical micelle concentration (CMC; 0.001% to 1%, depending on the detergent).

The gel concentration has a dramatic effect on resolution and should be optimized in order to achieve the best separation and band sharpness. In general, increasing the %T will improve band sharpness.

Troubleshooting

Gel polymerization at acid pH can be problematic, and sodium sulfite is needed for efficient polymerization (Andrews, 1986). Both the ammonium persulfate and the sodium sulfite must be freshly made, and the highest quality reagents available should be used. Furthermore, the gel solutions should be at room temperature for effective polymerization.

If the protein does not enter the gel and no stained material is present at the well surface, try reversing the polarity of the electrode. If material concentrates at the top of the gel, try lowering the acrylamide concentration. Stained material at the top of the gel may also indicate poor solubilization, and increasing the ionic strength of the solubilization buffer or adding a small amount of urea and/or nonionic detergent may be required.

Anticipated Results

Proteins will resolve depending on their solubility and native charge at the chosen pH. Ideally, a distinct band representing the protein of interest will be visible. If the band is diffuse, then increasing the gel concentration or using a gradient gel will improve resolution. If the band is not visible, then the protein may be at its isoelectric point or may have moved out of the gel because it had the wrong charge. Continuous gel systems, although more versatile, will give lower resolution than discontinuous gels. Detergents or other solubilizing agents such as urea may be needed to fully solubilize and resolve the protein. Once the conditions that resolve the protein are determined, Ferguson plots will give indications of multiple subunit structure, native size, and potential isoform relationships.

Time Considerations

Separations will be complete when the tracking dye or protein reaches the bottom of the gel. For minigels, this generally takes 1 to 2 hr, using 15 or 30 mA for 0.75- or 1.5-mm-thick gels, respectively. Standard-format gels require 4 to 5 hr at 15 or 30 mA for 0.75- or 1.5-mm-thick gels, respectively. Standard-format gels can also run overnight at 4 to 6 or 8 to 12 mA for 0.75- or 1.5-mm-thick gels, respectively.

Literature Cited

Andrews, A.T. 1986. Electrophoresis: Theory, Techniques and Biochemical and Clinical Applications, 2nd ed. Oxford University Press, New York.

Ferguson, K.A. 1964. Starch-gel electrophoresis—application to the classification of pituitary proteins and polypeptides. *Metabolism* 13:985-1002.

Hames, D. 1990. One-dimensional polyacrylamide gel electrophoresis. *In* Gel Electrophoresis of Proteins: A Practical Approach, 2nd ed. (B.D. Hames and D. Rickwood, eds.) pp. 1-147. Oxford University Press, New York.

Hedrick, J.L. and Smith, A.J. 1968. Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch. Biochem. Biophys.* 126:155-164.

Rodbard, D. and Chrambach, A. 1971. Estimation of molecular radius, free mobility, and valence using polyacrylamide gel electrophoresis. *Anal. Biochem.* 40:95-134.

Schägger, H. 1994. Native gel electrophoresis. *In* A Practical Guide to Membrane Protein Purification (G. Von Jagow and H. Schägger, eds.) pp. 81-104. Academic Press, San Diego.

Sigma. 1986. Nondenatured protein molecular weight marker kit (Technical Bulletin No. MKR-137). Sigma Chemical Company, St. Louis, Mo.

Key Reference

Andrews, 1986. See above.

Covers a variety of electrophoretic techniques, including nondenaturing electrophoresis and Ferguson plots.

Contributed by Sean R. Gallagher
Hoefer Pharmacia Biotech
San Francisco, California

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Two-Dimensional Gel Electrophoresis

Two-dimensional gel electrophoresis combines two different electrophoretic separating techniques in perpendicular directions to provide a much greater separation of complex protein mixtures than either of the individual procedures. The most common two-dimensional technique uses isoelectrofocusing (IEF) in a tube gel (see Basic Protocol 1) followed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) in a perpendicular direction (Basic Protocol 3). This combination of isoelectric point (pI) and size separation is the most powerful tool for protein separations currently available. After staining, proteins appear on the final two-dimensional gel as round or elliptical spots instead of the rectangular bands observed on one-dimensional gels. Although the total separating power of large-format two-dimensional gels is estimated to be as high as 5000 spots per gel, in practice a single two-dimensional separation of a complex mixture such as a whole-cell or tissue extract may produce 1000 to 2000 well-resolved spots when a sensitive detection method is used.

The most common IEF procedures are based on the use of soluble ampholytes, which are relatively small organic molecules with various isoelectric points and buffering capacities. The pH gradient for IEF gels is produced when the soluble ampholytes migrate in the gel matrix until they reach their isoelectric point. Because stable pH gradients outside the pH 3.0 to 8.0 range are difficult to create, alternative protocols using nonequilibrium conditions are required to resolve proteins with pI values below 3.0 to 4.0 (see Alternate Protocol 1 for acidic proteins) or above 8.0 (see Alternate Protocol 2 for basic proteins). One of the more important limitations of soluble ampholytes is the difficulty in obtaining highly reproducible pH profiles, especially when very narrow pH ranges are needed.

An increasingly attractive alternative to soluble ampholytes is the use of immobilized pH gradient (IPG) gels (see Basic Protocol 2). In this system, the buffering side chains are covalently incorporated into the acrylamide matrix, and any pH range and curve shape can be generated by pouring a gradient gel using two solutions that differ in ampholyte composition rather than acrylamide concentration. As with tube gels, the initial electrophoresis is followed by a second separation using SDS-PAGE in a perpendicular direction (see Basic Protocol 4). The use of IPG gels has recently increased, for at least three major reasons: many of the technical problems associated with their use have been solved or substantially minimized, reproducible premade IPG gels are now commercially available, and lately strong interest has arisen in using two-dimensional gels for proteome analysis studies (analyzing and comparing the complete protein profiles of cell lines, tissue samples, or single-celled organisms).

Another common two-dimensional electrophoresis format is a nonreducing/reducing electrophoretic separation (see Alternate Protocol 3), which provides useful information about intersubunit disulfides or protein-protein complexes that have been cross-linked using a bifunctional chemical cross-linker containing a disulfide bond within the linker region.

This unit also includes support protocols describing pI standards and pH profile measurements (see Support Protocol 1), casting Immobiline gels (see Support Protocol 3), preparation of tissue culture cells and solid tissues for isoelectricfocusing (see Support Protocols 4 and 5), preparation of molecular weight standards for two-dimensional gels (see Support Protocol 6), and two-dimensional protein databases (see Support Protocol 7).

NOTE: High-purity water (e.g., Milli-Q water or equivalent) is essential for all solutions. For cautions relating to electricity and electrophoresis, see Safety Considerations in the introduction to *UNIT 10.1*.

Contributed by Sandra Harper, Jacek Mozdzanowski, and David Speicher

Current Protocols in Protein Science (1998) 10.4.1-10.4.36

Copyright © 1998 by John Wiley & Sons, Inc.

BU MASI CENTRAL UNIVERSITY LIBRARY

HIGH-RESOLUTION EQUILIBRIUM ISOELECTROFOCUSING IN TUBE GELS

This protocol describes the preparation of broad-range first-dimension gels using soluble ampholytes that resolve proteins with pI values between approximately 3.0 and 8.0, and is based on the original procedure described by O'Farrell (1975). The procedure presented here refers specifically to 3-mm IEF tube gels (first-dimension) combined with 1.5-mm-thick 16 × 16–cm (size of separating gel) second-dimension gels (see Basic Protocol 3) and may be easily adapted to a variety of different gel sizes (see Table 10.4.1). A 3-mm IEF gel has a total protein capacity of ~500 µg for complex protein mixtures such as whole-cell extracts. The maximum capacity of any single protein spot is ~0.5 to 5 µg, depending on the solubility of the protein near its isoelectric point and the separation distance from any near neighbors.

In this protocol, gels are cast and prefocused before the sample is loaded. The proteins are then separated according to isoelectric point, and the gels are extruded from the tubes and stored. Measuring pH profiles in IEF gels is a convenient and accurate method for determining pI (see Support Protocol 1). To provide optimal reproducibility, multiple gels should be cast and run simultaneously. This is especially important for comparative studies involving complex mixtures of proteins.

The IEF gels may be cast either by pouring the gel solution into the gel tubes (steps 3a to 7a) or by using hydrostatic pressure (steps 3b to 7b). Pouring the gel solution into the gel tubes is convenient for 3-mm-diameter IEF gels and requires only a minimal excess of reagents. Because the gels are cast using a long needle and syringe, for narrower gels, where the needle does not fit inside the gel tube, casting using hydrostatic pressure is more appropriate. This method requires a larger excess of reagents and special casting cylinders. Many types of ampholytes are readily available from different suppliers to form the desired pH profiles. As ampholytes may vary significantly in their performance, careful selection of the appropriate ampholytes is usually necessary (see Commentary).

Materials

Chromic acid, in acid-resistant container
Urea (ultrapure)
30% acrylamide/0.8% bisacrylamide (see recipe)
20% (w/v) Triton X-100 (see recipe)
Ampholytes (e.g., pH 3-10/2D; ESA)
TEMED (*N*,*N*,*N*′,*N*′-tetramethylethylenediamine)
2.5% (w/v) ammonium persulfate (see recipe; prepare immediately before use)
8 M urea (see recipe; prepare immediately before use)
0.1 M orthophosphoric acid (H_3PO_4; see recipe)
0.1 M NaOH (*APPENDIX 2E*; make fresh daily)
Lysis buffer (see recipe)
Protein samples to be analyzed
Equilibration buffer (see recipe)
2-Mercaptoethanol

Isoelectrofocusing apparatus (e.g., Protean II xi 2D from Bio-Rad or equivalent) with glass tubes, casting stand, buffer chambers, rubber grommets, and plugs
37°C water bath
110°C oven
10-ml syringe equipped with filter capsule (0.22 or 0.45 µm, e.g., Costar µStar LB)
10-ml syringe equipped with blunt needle [e.g., 20-G × 6 in. (15 cm) or 18-G × 6 in. (15 cm)]

BCU IASL - CENTRAL UNIVERSITY LIBRARY

Large glass cylinder sealed at bottom with Parafilm (optional, for hydrostatic pressure casting method only)
2000-V power supply
60-ml syringe
Metal or plastic scoop
Dry ice pellets

Wash tubes and prepare the gel mixture

1. Remove the glass tubes from a chromic acid–filled container. Extensively wash the tubes with water, using high-purity water for the last wash. Dry the tubes at least 1 hr in an oven at 110°C and store them at room temperature, covered with aluminum foil.

   To prevent gels from sticking to the glass tubes, gel tubes have to be very clean. Satisfactory results are obtained by storing the tubes in chromic acid between uses and washing them shortly before use. Because drying the tubes requires at least 1 hr, cleaning steps should be performed the day before gels will be cast.

   CAUTION: *Chromic acid is highly corrosive; follow supplier's precautions carefully.*

2. Prepare the gel solution by mixing:

   16.9 g urea
   4.0 ml of 30% acrylamide/0.8% bisacrylamide
   3.0 ml of 20% (w/v) Triton X-100
   7.5 ml water
   3.0 ml ampholytes.

   Briefly warm the mixture in a 37°C water bath to solubilize urea if needed.

   To minimize decomposition of urea, never warm any solutions containing urea above 37°C, use ultrapure urea, and prepare solutions immediately before use.

   Choice of ampholyte composition is one of the key factors determining the quality of isoelectrofocusing separations. Substantial differences in performance, resolution, and shape of the pH gradient formed may be observed with different combinations of ampholytes and with ampholytes from different suppliers. ESA's ampholytes (pH 3-10/2D) are suited for most applications and give reproducible results.

   Although purity of all reagents is important, the purity of urea and choice of ampholytes are among the most critical factors for the quality and performance of isoelectrofocusing. Most commercially available reagents marketed specifically for two-dimensional gel electrophoresis should be suitable, although individual lots of reagents from any supplier may provide variability and/or unacceptable results.

Cast gels by pouring

3a. Wrap one end of each glass tube with Parafilm and mount the tube in a casting stand. Mark all the tubes to indicate the desired gel height.

   For reproducible results, all gels should be the same height.

4a. Filter the gel solution using a 10-ml syringe equipped with a syringe-tip filter capsule. Briefly degas the gel solution (~5 min) either by sonication or under vacuum. Then add 42.5 µl TEMED and 187.5 µl of 2.5% (w/v) ammonium persulfate solution to the filtered gel mixture and swirl gently to mix.

5a. Using a 10-ml syringe with a blunt needle, fill each glass tube with gel solution to the desired height. Make sure there are no air bubbles trapped in the gel.

   A needle is the best choice for casting gels if tubes of 3-mm inner diameter are used. For narrower tubes, the use of hydrostatic pressure is more appropriate (see steps 3b to 7b,

BOSTON CENTRAL UNIVERSITY LIBRARY

below). For long gels the needle can be extended by inserting a piece of capillary polyethylene tubing over the needle tip. The amount of gel solution described in step 2 is sufficient for sixteen 3-mm tube gels that are 16 cm long.

6a. Immediately overlay each gel with ~50 µl of 8 M urea.

A pipettor with a capillary pipet tip is a convenient tool for overlaying with urea. Avoid mixing the overlay and gel solutions. Polymerization starts to occur ~15 min after the addition of TEMED and ammonium persulfate. It is essential that the gels be poured and overlaid before significant polymerization has occurred.

7a. Let the gels polymerize at least 3 hr prior to use.

Urea decomposes at a substantial rate at room temperature; therefore, the gels should be used the same day they are cast.

Cast gels using hydrostatic pressure

3b. Place a rubber band around the gel tubes so they form a tight bundle. Place the bundle inside a larger glass cylinder that is sealed at the bottom with several layers of Parafilm. All tubes must be precisely vertical.

The dimensions of the larger cylinder depend on the dimensions and number of gel tubes. Excessive space will require more gel solution to cast the gels.

4b. Filter the gel solution using a 10-ml syringe and filter capsule. Degas the gel solution briefly (~5 min) either with sonication or under vacuum. Add 42.5 µl TEMED and 187.5 µl of 2.5% ammonium persulfate solution and swirl.

5b. Pipet the gel solution into the bottom of the glass cylinder. Gently run water down the outside of the tube bundle using a wash bottle. Keep adding water until the gel mix reaches the desired height.

Hydrostatic pressure will force the gel solution into the tubes. Sufficient gel solution must be used to obtain the desired gel height while avoiding forcing any water into the tubes. The volume of gel solution required can be estimated as follows: number of gels × 3.14 × (tube internal radius in cm)^2 × height in cm + ~10 ml to keep a safe level of gel mix at the bottom of the casting cylinder. As water is less dense than the gel solution, the water level will be slightly higher than the level of gel solution inside the tubes.

6b. Overlay the gels with 8 M urea.

Urea decomposes at a substantial rate at room temperature; therefore, the gels should be used the same day they are cast.

7b. Let the gels polymerize at least 3 hr prior to use.

Mount the gels in the electrophoresis unit

8. Prepare the lower electrode solution by degassing the proper amount of 0.1 M H_3PO_4 under vacuum with stirring for at least 5 min. Fill the bottom electrophoresis chamber.

The amount of phosphoric acid depends on the length of the gel tubes and the type of electrophoresis unit. The solution should cover the entire gel for good heat dissipation. Approximately 3 liters are required for Protean II xi 2D electrophoresis units.

9. Remove the gel tubes from the casting stand, remove the Parafilm from the tube bottoms, and inspect gels for irregularities or trapped air bubbles. Discard imperfect gels. If using gels cast with hydrostatic pressure, remove the bundle of tubes en bloc, cut off excess acrylamide with a razor blade, and then rinse away remaining acrylamide particles from the outside of each tube.

10. Place a rubber grommet on the top of the tube. Approximately 5 mm of the tube should be visible above the upper edge of the grommet.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

11. Mount the tube with the grommet in the upper reservoir and plug any unused holes.

After the tube is seated, its lower end must be submerged in the lower electrode solution. Be sure to remove any air bubbles trapped at the bottom of the tube by shaking or tapping the tube gently. Alternatively, with some units bubbles can be dislodged by raising and lowering the tubes or by using a long curved needle and syringe.

Prefocus the gels

12. Prepare the 0.1 M NaOH upper electrode solution by degassing under vacuum with stirring for at least 5 min.

The amount of upper electrode solution necessary depends on the type of electrophoresis chamber. If a Bio-Rad Protean II xi 2D apparatus is used, 1 liter of 0.1 M NaOH is sufficient for both prefocusing and the separation.

13. Remove the 8 M urea overlay from the top of the gels using a Pasteur pipet and place ~50 µl lysis buffer on the top of each gel.

14. Overlay lysis buffer with the degassed 0.1 M NaOH to fill the gel tubes. Avoid mixing of NaOH with the lysis buffer.

15. Pour the degassed 0.1 M NaOH into the upper chamber, making sure that all the gel tubes are covered with the electrode solution. Check carefully for leaks and air bubbles, then place lid on apparatus.

16. Connect the electrodes to a power supply by the red (+) lead to the lower chamber and the black (–) lead to the upper chamber.

The voltages and currents used during electrophoresis are dangerous and potentially lethal. Safety considerations are given in the Electricity and Electrophoresis section of UNIT 10.1.

17. Prefocus for 30 min using 500 V constant voltage.

Load the samples

18. Turn off power supply (see Safety Considerations in *UNIT 10.1*), disconnect leads, and remove lid. Using a 60-ml syringe, remove the electrode solution (0.1 M NaOH) from the upper chamber.

19. Remove the electrode solution and the overlay solution from each tube. Be careful not to damage the gel surface.

20. Place ~50 µl lysis buffer on the top of each gel. Wait at least 2 min.

21. Remove the lysis buffer from the tubes.

Rinsing the gels with lysis buffer removes any residual NaOH and protects the samples against exposure to high pH.

22. Load protein samples to be analyzed and carefully overlay each sample with ~50 µl lysis buffer diluted with water 8:2 (v/v). Avoid mixing the buffer with the sample.

The overlay solution protects samples from direct contact with the strong base used as an upper electrode solution. Dilution of the lysis buffer with water is necessary to decrease the density so the overlay does not mix with the sample.

A 3-mm-i.d. × 16-cm-long IEF gel has a total protein capacity of ~500 µg for whole-cell extracts and other complex protein mixtures. The maximum capacity for any single protein spot is ~0.5 to 5 µg, depending on its solubility near its isoelectric point and the separation distance from any near neighbors. Preparation of relatively pure protein samples for isoelectrofocusing is generally straightforward. The sample usually may be prepared in one of the following ways: dialyze into any compatible low-ionic-strength buffer, lyophilize in a volatile or compatible low-ionic-strength buffer and dissolve in lysis buffer, or

BOLIVAR CENTRAL UNIVERSITY LIBRARY

precipitate the protein using trichloroacetic acid (TCA) and redissolve in lysis buffer. For preparing extracts from cultured cells and from tissue samples, see Support Protocol 4 and Support Protocol 5, respectively.

The minimum sample concentration of protein or radioactivity has to be sufficient for the desired detection method. For complex protein mixtures such as tissue or cell extracts, a 500-µg total load is recommended for Coomassie blue staining (UNIT 10.5) or electroblotting (UNIT 10.7) for subsequent structural analysis, a 50-µg total protein load should be sufficient for silver staining (UNIT 10.5) or immunoblotting, and no less than 100,000 counts/gel is recommended for proteins labeled with ^{3}H, ^{14}C, or ^{35}S for autoradiography purposes. Sample volumes should be <150 µl for 3-mm gels and <40 µl for 1.5-mm gels. This implies at least a 5 µg/µl protein concentration in the sample for gels to be stained with Coomassie blue.

23. Carefully fill all tubes with 0.1 M NaOH. Avoid mixing the NaOH solution with the overlay solution and the sample.

24. Fill the upper reservoir with 0.1 M NaOH. Be sure that all gel tubes are covered with the solution.

Run the gels

25. Connect the electrodes to a power supply with red (+) to the lower chamber and black (–) to the upper chamber.

26. Focus for a total of 12,000 Vhr.

Unlike other electrophoretic techniques, in IEF the volt-hour is the most common unit describing the "time" of isoelectrofocusing. The initial voltage is usually set according to the desired number of volt-hours in a way that is convenient for the operator (i.e., so that the separation will run overnight), but it should not be <400 V. The upper voltage limit is restricted by heat released in the gels during isoelectrofocusing. At constant voltage the current will be the highest during the first hour of separation. The initial current will be strongly influenced by the ionic strength of the samples loaded onto the gels. An initial voltage of <800 V is recommended for 3-mm gels loaded with samples containing less than 100 mM salts/buffers; the voltage could be increased to 1200 V after ~1 hr, if cooling is used. The current is a derivative of voltage and is never preset for isoelectrofocusing purposes. Some power supplies allow preprogramming the desired number of volt-hours and continuously adjust voltage and current during the isoelectrofocusing procedure (constant power). The total number of volt-hours is a major factor that affects separation in the first dimension. Optimal focusing time will vary for different ampholyte combinations, but 12,000 Vhr is a reasonable value for most systems. To achieve a total of 12,000 Vhr set the power supply to 667 V for 18 hr. These conditions are convenient for an overnight separation and do not require use of a cooling unit. Higher voltages can be used but may cause overheating of gels unless a highly efficient cooling system is employed. The maximum practical voltage decreases with increased gel tube inner diameter. Focusing for too long may cause cathodic drift and result in a shifted pH profile in the gel, whereas focusing for a short time will decrease resolution.

Extrude and store gels

27. Turn off power supply and carefully disconnect leads. Detach the lid and remove the NaOH solution from the upper reservoir of the electrophoresis chamber using a 60-ml disposable plastic syringe.

28. Remove one gel tube at a time from the chamber.

29. Using a 10-ml syringe equipped with a blunt needle, slowly and carefully inject water between the gel and glass tube. Start from the bottom of the tube, then repeat the procedure from the top. The gel should slide out of the glass tube.

It is convenient to let the gel slide from the glass tube onto a metal or plastic scoop, which facilitates transfer of the gel into a storage vial. It is relatively easy to break the gel during

BCU IAS / CENTRAL UNIVERSITY LIBRARY

extrusion, and practicing on several unused gels is recommended. To extrude smaller-diameter gels, use water pressure generated by a syringe connected to the gel tube with Tygon tubing. If clean, unscratched glass tubes are used, extrusion should be easy.

30. Using the scoop, slide the gel into a 4.5-ml cryovial containing 3 ml equilibration buffer and 50 µl 2-mercaptoethanol. Close the vial, incubate exactly 5 min at room temperature, then freeze by placing the tube horizontally on top of dry ice pellets. Do not move or agitate the tube while the sample is freezing.

The IEF gels may be run on a second-dimension gel immediately (see Basic Protocol 3), or can be stored at −80°C for many weeks. Even when the second-dimension is to be run immediately, extruded gels should be frozen after a carefully controlled incubation time at room temperature, such as the 5 min cited above for 3-mm-i.d. gels, to minimize diffusion of proteins out of the IEF gel. This short incubation before freezing will allow glycerol to diffuse into the gel. Too short an incubation or agitation during freezing can result in gel breakage. The total incubation time in equilibration buffer (sum of the time prior to freezing and after thawing) is critical and should be carefully controlled. Insufficient incubation time in equilibration buffer will not allow sufficient time for SDS to diffuse into the gel and saturate sites on the proteins. Excessive incubation times can result in appreciable protein losses due to diffusion out of the highly porous IEF gel.

CONDUCTING pH PROFILE MEASUREMENTS

SUPPORT PROTOCOL 1

Standards with different isoelectric points can help in evaluating the performance of a specific system and determining the effective pH range in the isoelectrofocusing gel. Many pI standards are commercially available from different suppliers. It is most useful to separate a mixture of standard proteins that is prepared from several individual proteins or purchased as a preformulated kit. This mixture should be run in parallel with experimental samples on a separate reference gel. It is generally not recommended to run pI standards together in the same gel with samples because of possible interference with migration and identification of proteins of interest. Instead of analyzing standard proteins, a more precise evaluation of the pH profile can be made by directly measuring the pH throughout the gel using either a surface pH electrode or the following procedure.

1. Prepare and focus one or two gels (see Basic Protocol 1, steps 1 to 26) without any sample in parallel with experimental samples.
2. Prepare 20 to 40 glass test tubes each containing 1 ml high-purity, degassed water for each gel that will be used to measure the pH gradient (measurements on duplicate gels are recommended).

   The number of tubes required per gel equals twice the gel length (in cm).

3. After electrofocusing is completed, extrude the blank gels (see Basic Protocol 1, steps 27 to 29). Briefly rinse the gels with water.

   After extrusion, gel surfaces may be contaminated with electrode solutions. Rinsing with water is essential for obtaining reliable pH profiles.

4. Place the gel on a glass plate with a plastic ruler below the plate. Cut the gel into 0.5-cm pieces using a sharp razor blade.
5. Place each gel piece in a test tube containing 1 ml water.

   Do not mix the order of samples because each gel piece represents a single pH profile data point.

6. Place all test tubes on a shaker and shake gently for 1 hr at room temperature.
7. Read the pH of each solution and plot the pH profile as a function of the distance from the top of the gel.

BOC IASI CENTRAL UNIVERSITY LIBRARY

ALTERNATE PROTOCOL 1

NONEQUILIBRIUM ISOELECTROFOCUSING OF VERY ACIDIC PROTEINS

Basic Protocol 1 is sufficient for separating proteins with isoelectric points greater than ~3.0 to 3.5. For very acidic proteins, however, a nonequilibrium system is needed. The major features of this method are utilization of a shorter focusing time (without reaching equilibrium), a modified ampholyte mixture, and different electrode solutions.

Additional Materials (also see Basic Protocol 1)

10% (w/v) ammonium persulfate (prepare immediately before use)
Concentrated sulfuric acid (used in lower chamber electrode solution)
Ampholytes, pH 2-11 (used in upper chamber electrode solution)

To analyze very acidic proteins, follow Basic Protocol 1 with these exceptions in the indicated steps:

2. When preparing the gel solution, use the following mixture of ampholytes: 2.4 ml ampholytes pH 2.5-4 and 0.6 ml ampholytes pH 2-11.

4. Following the procedure for casting gels by pouring, add 100 µl of 10% ammonium persulfate solution, swirl, add 42.5 µl TEMED, and swirl again.

   Gel mixtures containing entirely or predominantly very acidic or very basic ampholytes are generally difficult to polymerize. Use of an increased ammonium persulfate concentration and adherence to the proper order of adding the reagents should ensure polymerization.

8. Prepare the bottom chamber electrode solution by adding 4.5 ml concentrated sulfuric acid to 3 liters water. Degas at least 5 min.

Omit steps 12 to 19 (do not prefocus the gels).

20. Remove the 8 M urea (polymerization overlay solution) and place ~50 µl lysis buffer on top of each gel. Wait at least 2 min, then remove the lysis buffer.

23. Carefully fill all tubes with the upper chamber electrode (anode) solution prepared by mixing pH 2-11 ampholytes with water in a 1:40 ratio.

24. Fill the upper buffer chamber (anode) with the solution described in step 23.

    Iminodiacetic acid (10 mM) may be a more economical alternative anode solution.

26. Focus for a total of 4000 Vhr.

ALTERNATE PROTOCOL 2

NONEQUILIBRIUM ISOELECTROFOCUSING OF BASIC PROTEINS

In general, most equilibrium IEF gel systems using soluble ampholytes produce pH gradients that do not exceed pH 8.0 on the basic end, yet many proteins have higher pI values. For this reason samples containing very basic proteins are usually focused using a nonequilibrium system. In an equilibrium system, proteins are loaded on the basic end of the gel and migrate toward the acidic end until they reach a pH equal to their pI. In nonequilibrium systems, the sample is loaded on the acidic end of the gel, and focusing is terminated after a relatively short time (fewer volt-hours).

To run nonequilibrium IEF gels, follow the procedure previously described (see Basic Protocol 1) with these alterations in the indicated steps:

8. Use 0.1 M NaOH as the lower electrode solution.

   Electrode solutions and electrodes are reversed in this procedure relative to equilibrium isoelectrofocusing.

BCU IAS CENTRAL UNIVERSITY LIBRARY

Omit steps 12 to 19 (do not prefocus the gels).

20. Remove the 8 M urea (polymerization overlay solution) and place 50 µl lysis buffer on top of each gel. Wait at least 2 min, then remove the lysis buffer.

23. After loading the samples and overlaying with lysis buffer diluted with water 8:2 (v/v) as in Basic Protocol 1, use 0.1 M H_3PO_4 instead of NaOH to fill all gel tubes.

24. Use 0.1 M H_3PO_4 as the upper electrode solution.

25. Reverse the connection of electrodes—i.e., connect the red (+) lead to the upper chamber and the black (–) lead to the lower chamber.

26. Focus for a total of 3000 to 5000 Vhr.

The optimal number of volt-hours depends on the nature of the sample and the ampholytes used. The values recommended above may need to be adjusted empirically.

ISOELECTROFOCUSING USING IMMOBILIZED pH GRADIENT GEL STRIPS

BASIC PROTOCOL 2

In immobilized pH gradient (IPG) gels, the ampholytes are covalently linked to the acrylamide matrix, which facilitates production of highly reproducible gradients as well as very narrow pH gradients for optimal resolution of minor charge differences. A variety of precast gels and all the necessary equipment are commercially available from Hoefer Pharmacia. Equipment and chemicals are also available for the user to cast gels in the laboratory (see Support Protocol 3), although precast gels are likely to suffice for the majority of applications. Narrow strips of precast IEF gels (Immobiline DryStrips) may be used to achieve a first-dimension separation for two-dimensional gel electrophoresis, and broader precast slab gels (Immobiline DryPlates) can be used to compare multiple samples after IEF separation only (see Support Protocol 2 and Table 10.4.3). In this protocol, precast Immobiline DryStrips from Hoefer Pharmacia are rehydrated overnight using the reswelling cassette (one to twelve sample strips may be handled at a time); samples are applied using sample cup holders and gel strips are isoelectrofocused overnight. This procedure has been adapted from instruction booklets provided by Hoefer Pharmacia with Immobiline Dry Strip Kits and with the Immobiline DryStrip reswelling tray. Since there is a greater selection of pH ranges for premade Immobiline DryPlates than DryStrips, it is sometimes convenient to cut DryPlates into strips prior to rehydrating the gel to obtain narrower pH ranges where needed. See Basic Protocol 4 for details concerning preparing and running the second-dimension gel.

Wear gloves throughout the procedure and handle the Immobiline DryStrips with forceps where feasible to prevent extraneous protein contamination of the gels and gel solutions. Thoroughly clean all equipment with a mild laboratory detergent solution, rinse well with Milli-Q water, and allow to dry before using. Solutions containing 10 M urea may be heated *briefly* to 30° to 40°C to aid in solubilization.

Materials

Urea (ultrapure)
CHAPS or Triton X-100
Pharmalyte 3-10, 4-6.5, and/or 8-10.5 soluble ampholytes (see Table 10.4.1; Hoefer Pharmacia)
Ampholine pH 6-8 (Hoefer Pharmacia)
DTT (dithiothreitol)
Bromphenol blue
Precast Immobiline DryStrips (Hoefer Pharmacia)

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Table 10.4.1 Rehydration Solutions for Immobiline DryStrips[a]

Component	Final conc.	DryStrip type		
		3-10L	3-10NL	4-7L
Ultrapure urea	9 M[b]	2.7 g	2.7 g	2.7 g
CHAPS[c]	2%	0.1 g	0.1 g	0.1 g
Pharmalyte pH 3-10	1:50	100 µl		
Pharmalyte pH 4-6.5			50 µl	100 µl
Pharmalyte pH 8-10.5			25 µl	
Ampholine pH 6-8			25 µl	
DTT	0.3%	75 mg	75 mg	75 mg
Bromphenol blue	Trace	A few grains	A few grains	A few grains
Milli-Q water		To 5 ml	To 5 ml	To 5 ml

[a]Rehydration solutions should be prepared fresh immediately before use and should be filtered using a 0.2-µm filter. Minimize total time the solution is at room temperature prior to use to minimize decomposition of urea. If the reswelling tray is used, ~250 or 400 µl rehydration solution is required per 11- or 18-cm DryStrip, respectively.

[b]Urea concentrations between 8 M and 10 M are typically used. Higher concentrations promote better protein solubility during isofocusing while increasing the risk of urea crystallization. If 10 M urea is used, extra care should be taken to minimize evaporation and the temperature should be maintained at 20°C or slightly higher at all times to avoid crystallization of the urea.

[c]The optimal detergent and detergent concentration should be empirically determined. Other common alternatives are Triton X-100 and octyl-glucoside. The detergent used must be nonionic or zwitterionic to avoid high current and consequent overheating during isoelectrofocusing.

DryStrip cover fluid (Hoefer Pharmacia)
Immobiline DryStrip kit (Hoefer Pharmacia) including:
- Cathode electrode
- Anode electrode
- Sample cup bar
- Tray
- Sample cups
- Immobiline strip aligner
- IEF electrode strips
- Sample application pieces
- Instruction manual

Protein sample to be analyzed
Lysis buffer (see recipe)

Immobiline DryStrip reswelling tray (Hoefer Pharmacia)
Forceps
Filter paper
Glass plate
Flatbed electrophoresis unit (Hoefer Pharmacia Multiphor II or equivalent)
Recirculating cooling water bath
Power supply (minimum capacity of 3000 to 3500 V)
Petri dishes

Additional reagents and equipment for protein detection by staining (*UNIT 10.5*) and/or for electroblotting (*UNIT 10.7*, optional)

BCU IASI - CENTRAL UNIVERSITY LIBRARY

Rehydrate the Immobiline DryStrip(s)

1. Prepare an appropriate rehydration solution for the type of DryStrips to be used as described in Table 10.4.1 (~400 µl rehydration solution per 18-cm DryStrip).

   The urea concentration in the rehydration solution should be 8 to 10 M, and 2% CHAPS or another appropriate detergent (zwitterionic or nonionic) such as Triton X-100, NP-40, or n-octylglucoside should be included in the rehydration solution to aid in sample solubility. The optimal detergent and detergent concentration may vary with type of sample and should be determined empirically.

   One possible method of loading large sample volumes onto IPG gels is to add the sample directly to the rehydration solution. However, this sample loading method is not typically recommended since it tends to result in poor and variable protein loading into the gel. During rehydration of the gel, proteins are only poorly adsorbed into the gel matrix, apparently as a result of charge repulsion effects caused by the highly charged gel matrix.

   Solutions containing urea should be filtered using a 0.2-µm filter before use.

2. Slide the protective lid off the reswelling tray and level the tray by adjusting the leveling feet until the leveling bubble is centered.

3. For an 18-cm gel, pipet 350 to 400 µl of rehydration solution into a slot of the reswelling tray. Move the pipet along the length of the well while adding the solution to spread it evenly throughout the length of the slot. Avoid excessive air bubble formation while pipetting this solution.

4. Remove the protective cover from the Immobiline DryStrips and gently place them, gel side down, into the prepared slot.

   To facilitate their removal after rehydration, the strips should be oriented with their pointed ends at the sloped end of the slots in the rehydration tray. Be careful not to trap any air bubbles under the gel strips.

5. Overlay each strip with 2 to 3 ml of DryStrip cover fluid to prevent evaporation and urea crystallization. Slide the protective lid into place and allow gels to rehydrate overnight (~16 hours) at room temperature.

   Shorter rehydration times may not completely and reproducibly rehydrate the gels. Do not substantially exceed 16 hr as extensive incubation, especially rehydrating gels over a weekend, increases potential problems due to evaporation and subsequent urea crystallization. In addition, long incubation times increase the extent of urea decomposition, which will increase the risk of amino group modification on proteins by the cyanate produced from urea decomposition.

6. After the overnight rehydration, slide the lid off the reswelling tray. Place a forceps tip into the slight depression under each strip and remove the strip. Gently blot any excess oil or moisture from the plastic backing of the rehydrated strips with filter paper. A damp piece of filter paper may also be used to blot the surface of the gel. Some of the paper may adhere to the gel and should be gently peeled away. The gel is now ready to be placed in the strip aligner on the cooling plate of the electrophoresis unit.

   Do not allow the gel to dehydrate prior to placing it on the cooling plate in step 11. (Steps 7 to 10 should be completed prior to removing the strips from the reswelling tray.)

Run the first dimension

7. Level the Multiphor II electrophoresis unit, then connect it to a circulating cooling water bath. Allow it to cool to 15°C for 1 to 2 hr to ensure even cooling. Do not cool below 15°C to prevent precipitation of urea in the gels.

BOSTON UNIVERSITY MEDICAL LIBRARY

8. Pipet ~5 ml DryStrip cover fluid onto the surface of the Multiphor II cooling plate. Position the Immobiline DryStrip tray on the cooling plate oriented with the red (+, anodic) electrode at the top, near the cooling tubes.

Avoid large air bubbles between the cooling plate and the tray (small bubbles should not cause a problem).

9. Connect the red and black electrode leads on the tray to their respective positions on the Multiphor II unit. Pour 10 ml of DryStrip cover fluid into the tray. Place the Immobiline DryStrip aligner on top of the oil, groove side up.

Avoid getting oil on top of the strip aligner. The possible presence of small air bubbles under the strip aligner is not important.

10. Cut two electrode strips to a length of 11 cm (regardless of the number of DryStrips used). Place the electrode strips onto a clean glass plate and soak each one with 0.5 ml Milli-Q water. Blot with a Kimwipe or tissue paper to remove excess water.

The electrode strips should be evenly soaked and just damp after blotting. Excessive water could cause sample streaking.

11. Transfer the strips from step 6 to adjacent grooves in the aligner tray. Position the rounded (acidic) end of each strip near the top of the tray at the red electrode (anode) near the cooling tubes, and the square end at the bottom of the tray near the black electrode (cathode). Be sure that the edges of all gel strips at the anode end are lined up evenly.

12. Place the blotted electrode strips from step 9 on top of the gel surface of the DryStrips near the anode and cathode ends of the gel. Position the red (anode) and black (cathode) electrodes on top of the electrode strips at their respective ends.

After the electrodes have been pressed down on top of the electrode strips, check that the gel strips have not shifted position.

13. Push the sample cups onto the sample cup bar. Place the sample cup bar near the anode end of the gel so that the small spacer arm just touches the electrode and the sample cups are nearest to the electrode, but do not allow the cups to touch the gel.

The sample cups should face the nearest electrode. The acidic end of the gel can usually be used for sample application; however, the optimal loading position may need to be determined empirically for different types of samples. At high protein concentrations and/or at non-optimal pH, samples may precipitate in the gel at the loading position.

14. Position one sample cup above each gel strip and push down to ensure good contact between the bottom of the sample cup and the gel strip. Make sure the gel strips have not shifted position.

15. Pour 70 to 80 ml of DryStrip cover fluid into the tray (it will cover the gels). If oil leaks into the sample cups, adjust the cups to stop leakage. When there is no leakage into the sample cups, add enough cover fluid to the tray to completely cover the sample cups (~150 ml).

16. Pipet protein samples (in lysis buffer) into the sample cups by underlaying. The sample should sink to the bottom of the cup. Check for leakage of the sample out of the sample cup.

Samples should either be lyophilized and then solubilized in lysis buffer, or diluted 9 parts lysis buffer to 1 part sample. The maximum volume each sample cup holds is 100 µl. The complexity of the sample, the sample solubility at the loading concentration and pH used, the thickness of the second-dimension gel, and the detection method to be employed should be considered when deciding how much protein to load. As a starting reference, typical

BCU IASI CENTRAL UNIVERSITY LIBRARY

loading ranges for 1.0- to 1.5-mm-thick 18-cm × 18-cm gels would be ~5 to 20 ng per major spot for silver staining and ~1 to 5 µg per major spot for Coomassie blue staining. When very complex samples are used such as whole cell extracts, total protein loads are likely to be ~20 to 100 µg for silver staining and ~200 to 1000 µg for Coomassie blue staining. The salt concentration in samples should be kept <50 mM and, if the sample contains SDS, the final SDS concentration should be <0.25%.

17. Place the lid on the Multiphor II unit and connect the leads to a power supply. Focus the gels with constant voltage for 2 to 3 hr at 500 V followed by 12 to 16 hr at 3500 V for a total of 40 to 60 kVhr. Refer to the user manual for exact recommended voltage conditions for each type of Immobiline DryStrip.

The optimal number of Vhr will depend upon the pH range of the Immobiline Strip used, the type of sample, and the sample load and volume; therefore, the optimal Vhr should be empirically determined for different applications.

18. When isoelectrofocusing is complete, disconnect the power supply and remove the cover from the Multiphor II unit. Remove the electrodes, electrode strips, and sample cup bar from the tray.

If gels are to be run in the second dimension immediately after isofocusing, steps 1 to 3 of Basic Protocol 4 should be completed prior to terminating isofocusing.

19. Using forceps, remove the DryStrips from the tray. If the gels are to be run in the second dimension immediately, place in a petri dish with the support film along the wall of the dish and proceed directly to equilibration of the gel (see Basic Protocol 4, step 4). Alternatively, gels may be stored sealed in a plastic bag at −80°C until ready to run the second-dimension gel.

Gels may be stored at least 2 to 3 months at −80°C. Do not place in the equilibration buffers required for the second dimension prior to storage.

ELECTROPHORESIS ON IMMOBILIZED pH GRADIENT GELS

SUPPORT PROTOCOL 2

In this protocol, after precast IEF gels (Immobiline DryPlates) from Hoefer Pharmacia are rehydrated, samples are loaded and subjected to isoelectrofocusing. Gels are typically run at 2500 to 3500 V and require focusing times of 2 to 7 hr. Protein samples may be detected by conventional methods such as Coomassie blue or silver staining (*UNIT 10.5*). Isoelectric points can be determined with the use of pI calibration proteins; alternatively, because the gradient is linear, one can measure the migration distance across the gel and estimate the pI at each location. As noted in Basic Protocol 2, Immobiline DryPlates can be cut into 3-mm-wide strips to use as the first dimension of two-dimensional gels, since DryPlates are available in narrower pH ranges than DryStrips. DryPlates can also be used as described in this protocol to simultaneously separate multiple samples in a single dimension. Applications of this method include initial screening of samples to determine the optimal pH gradient prior to running more time-consuming two-dimensional gels, prescreening fractions from a chromatographic purification step prior to running two-dimensional gels, and evaluation of charge heterogeneity of purified proteins.

Additional Materials (also see Basic Protocol 2)

Precast DryPlate gel (Hoefer Pharmacia)
Repel-Silane (Hoefer Pharmacia)
Paraffin oil
Protein samples to be analyzed
Reswelling Cassette kit (Hoefer Pharmacia) including:
125 × 260 × 3–mm glass plate with 0.5-mm U frame
125 × 260 × 3–mm glass plate

continued

BOSTON UNIVERSITY LIBRARY

Silicone tubing
Pinchcock
Clamps
20-ml syringe
Roller (Hoefer Pharmacia)
Whatman no. 1 filter paper

Flatbed electrophoresis unit (Hoefer Pharmacia Multiphor II)
10° *or* 15°C cooling water bath
Electrode strips
Sample applicator strip or sample application pieces
Power supply (minimum capacity 3000 to 3500 V)

Additional reagents and equipment for protein detection by staining (*UNIT 10.5*) and for electroblotting (*UNIT 10.7*; optional)

Rehydrate the gel

1. Remove precast gel from packaging. If the entire gel is not needed, cut off the required number of lanes and reseal the unused gel. Mark the polarity of the gel section to be used by cutting a small triangle off the anode corner. Handle the gel by the support film only.

   It is critical that the lanes are cut from the gel in the proper orientation to preserve the pH gradient (see Fig. 10.4.1), and polarity must be indicated for proper orientation of electrodes later in the procedure.

2. Use the Reswelling Cassette to rehydrate the gel. Connect silicone tubing through hole in the bottom corner of the U-frame plate, seal with silicone glue, and connect the pinchcock to the other end of the tubing. Place a glass plate on a clean flat surface and wet with a few drops of water. Place the gel on the plate, gel side up. Gently roll with a clean rubber roller to remove any air bubbles.

3. Cover the plate and gel with the plate fitted with the U frame.

   The U-frame plate should be coated with a thin layer of Repel-Silane to prevent the gel from sticking to the plate.

4. Place clamps around the edges of the plates, making sure the seal is tight.

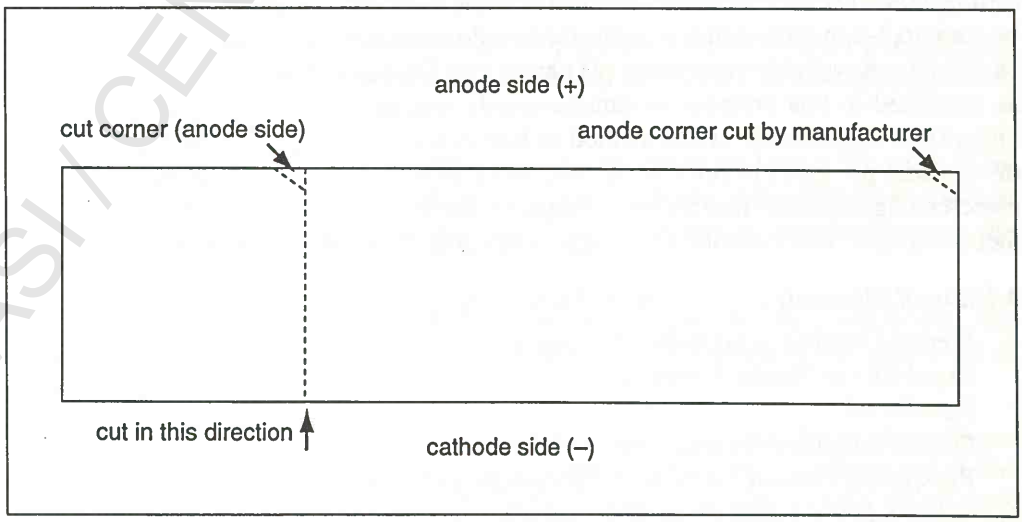


Figure 10.4.1 Marking orientation of a precast IPG gel when only a portion of the gel is used.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

5. Slowly fill the cassette with the desired rehydration solution using a 20-ml syringe connected to the silicone tubing and let stand for the recommended amount of time. Precool electrophoresis unit 1 to 2 hr prior to electrophoresis (see step 9).

   Reswelling with water for 2 to 3 hr is normally sufficient. If using additives such as urea, Triton, glycerol, or reducing agents, allow the gel to rehydrate overnight. Additives can be used to improve solubility of proteins near their isoelectric point. Reducing reagents such as 2-mercaptoethanol or DTT are used to reduce disulfide bonds.

6. When gel has been allowed to rehydrate completely, remove the clamps and gently pry the plates apart.

7. Moisten a piece of filter paper with water and place on top of the gel, then layer with a piece of dry filter paper.

8. Gently blot the gel by rolling over the dry filter paper with the rubber roller to remove excess water. The gel is now ready to be placed on the cooling plate.

   Do not let the gel dehydrate prior to placing it on the cooling plate in step 11.

Run the gel

9. Connect the flatbed electrophoresis unit to a recirculating cooling water bath. Allow to cool to 10°C for 1 to 2 hr to ensure even cooling. If the gel has been rehydrated in the presence of urea, do not cool below 15°C so that urea does not precipitate.

10. Pipet 2 to 3 ml paraffin oil onto the surface of the cooling plate.

11. Position the gel on the cooling plate, being careful not to trap air bubbles between the gel and the plate. Orient the gel so that the polarity of the gel matches the polarity of the cooling plate.

12. Soak two electrode strips with ~3 ml water, then blot to remove excess water.

13. Lay a blotted electrode strip along each long edge of the gel. Cut off the ends of the electrode strip so that it does not extend beyond the edge of the gel.

14. Load protein samples to be analyzed onto the gel. Use an applicator strip for sample volumes between 5 and 20 µl (make sure contact between the strip and the gel is uniform). Use sample application pieces for sample volumes >20 µl. Remove the application pieces halfway through focusing. For sample volumes of 2 to 10 µl, samples may be spotted directly on the gel without using applicator strips. See manufacturer's instructions for further details.

    An important experimental consideration is the position in the pH gradient where the sample is applied. The acidic end of the gel can usually be used for sample application; however, the optimal loading position may need to be determined empirically for different types of samples. At high protein concentrations and/or at nonoptimal pH's, samples may precipitate in the gel at the loading position.

    Samples should contain <50 mM salt or buffer components; greater concentrations will cause local overheating of the gel. If possible, salt-free samples should be solubilized or dialyzed in the rehydration buffer.

15. Align the electrodes with the electrode strips, put the safety lid in place, and connect the apparatus to the power supply. Conduct electrophoresis at 3000 V.

    Broad-range IPG gels, such as Immobiline DryPlate, pH 3-10, should be run at ~3000 V for 2 to 4 hr. Narrower-range gels are also run at 3000 V but may require a focusing time of 4 to 7 hr.

BOTANY LIBRARY UNIVERSITY OF FLORIDA

16. After removing gels from the electrophoresis apparatus, detect proteins using any conventional staining technique such as Coomassie blue or silver staining.

17. Preserve the gels by sealing in a plastic bag or by drying for a permanent record. Alternatively, electrotransfer the proteins on the gel to a membrane.

To dry a gel, presoak it first in a preservation solution. For silver-stained gels, use a solution of 5% to 10% (w/v) glycerol/30% (v/v) ethanol; for Coomassie blue–stained gels, use a solution of 5% to 10% (w/v) glycerol/16% (v/v) ethanol/8% (w/v) acetic acid. After soaking the gel, place it on a glass plate gel side up, cover with a cellophane sheet soaked in preservation solution, and allow to dry at room temperature.

For electrotransfer (UNIT 10.7), use film remover to remove the plastic support film from the gel. Electrotransfer of proteins to a polyvinylidene difluoride (PVDF) membrane using a Multiphor II NovaBlot transfer kit (Hoefer Pharmacia) is recommended. Transferring IPG gels requires special procedures; see the transfer kit manual for instructions.

SUPPORT PROTOCOL 3

CASTING AN IMMOBILINE GEL

An alternative to precast IPG gels is the use of Hoefer Pharmacia Immobilines to cast immobilized pH gradient gels with customized pH gradients and ranges, including very narrow pH ranges, to improve separation of proteins with small charge differences. This protocol describes the general procedure of casting custom-made Immobiline gels. The Reswelling Cassette used in Basic Protocol 2 for rehydrating gels is employed for casting the gels, which are polyacrylamide gels poured with a gradient of Immobilines following instructions provided by Hoefer Pharmacia application note 324.

Additional Materials (also see Support Protocol 2)

GelBond PAG film (Hoefer Pharmacia)
Immobiline solutions (Hoefer Pharmacia)
2.5% (v/v) glycerol

Gradient maker
Orbital shaker

Additional materials and equipment for rehydrating immobilized pH gradient gels (see Basic Protocol 2)

Cast the gel

1. Coat the plate with the U frame with Repel-Silane to prevent the gel from sticking to the glass plate.

2. Place a glass plate on a clean, flat surface and wet with several drops of water. Cover with a sheet of GelBond PAG film, hydrophilic side up. Use the roller to remove any air bubbles trapped between the film and the glass plate.

3. Place the plate with the U frame on top of the GelBond PAG film. Clamp the plates together on three sides.

4. Mix the Immobiline solutions following the instructions provided by Hoefer Pharmacia application note 324 to prepare the desired pH range. Cast the pH gradient gel using a gradient maker.

Once the catalysts have been added to the gel solution, it is important to work quickly to ensure that the gradient is poured before polymerization occurs. See UNIT 10.1 for instructions for using gradient makers.

5. Do not disturb the gel during the first 10 min to allow the gradient to stabilize. Allow the gel to polymerize 1 hr in a 50°C oven.

BCU IASI/ CENTRAL UNIVERSITY LIBRARY

Dry and store the gel

6. Allow the gel to cool to room temperature, then disassemble the cassette. Cut off a small corner to label the anode end of the gel.
7. Wash the gel 2 to 3 hr with 200 to 300 ml water. Use an orbital shaker and change the water two or three times.

   At this stage the gel may be used immediately (if no additives are needed) or dried as described in steps 8 to 10.
8. Wash the gel 30 min to 1 hr with 200 to 300 ml of 2.5% (v/v) glycerol.
9. Place the gel on a glass plate (gel side up) in a dust-free environment and allow to dry at room temperature overnight.
10. Store the dried gel in a sealed plastic bag at –20°C for up to 2 months.

    The gels may be rehydrated when needed (see Support Protocol 2).

PREPARING TISSUE CULTURE CELL EXTRACTS FOR ISOELECTROFOCUSING

SUPPORT PROTOCOL 4

Preparation of samples for isoelectrofocusing containing relatively pure proteins is generally straightforward (see Basic Protocol 1, step 22). In contrast, complex samples such as whole-cell extracts, tissue extracts, or subcellular fractions are more difficult to prepare for successful isoelectrofocusing. Solubility limitations both prior to isoelectrofocusing and during focusing restrict analysis of these complex samples to protocols that include nonionic detergents and urea (see Basic Protocol 1). In addition, the presence of DNA and RNA in crude cell extracts further complicates isoelectrofocusing. The protocol presented below is suitable for preparing samples from cell cultures and is based on quantities compatible with silver staining or Coomassie blue staining. If smaller cell numbers and high-sensitivity detection methods such as autoradiography are used, volumes and quantities should be adjusted as needed.

In this protocol the cells are harvested and washed in phosphate-buffered saline (PBS) with proteolysis inhibitors, then lysed in Tris/SDS buffer using sonication, after which the total protein concentration is determined in the lysate. The lysate is further treated with a mixture of DNase and RNase, and additional SDS and reducing agent are added. At this stage, the samples can be stored at –80°C for an extended time or, after addition of urea and lysis buffer, may be loaded directly onto prefocused IEF gels. Filtration of the final sample prior to loading onto the IEF gel is essential for quality of isoelectrofocusing.

Materials

Cell culture flasks containing cells of interest
PBS with proteolysis inhibitors (PBS/I buffer; see recipe)
Dry ice/ethanol (optional, for freezing samples)
Tris/SDS buffer (see recipe)
BCA protein assay kit (Pierce)
DNase and RNase solution (see recipe)
20% (w/v) SDS (*APPENDIX 2E*)
2-Mercaptoethanol
Urea (ultrapure)
Lysis buffer (see recipe)
50-ml centrifuge tube
Centrifuge with rotor (e.g., Beckman JS-4.2), 4°C

ROYAL CENTRAL UNIVERSITY LIBRARY

1- to 2-ml cryovials
Microcentrifuge, 4°C
Sonicator with microtip
0.2-µm microcentrifuge filter units (e.g., Millipore Ultrafree-MC filter units)

Harvest and wash the cells

1. Place cell culture flasks containing cells of interest on ice.
2. Rapidly wash cells three times with 2 to 6 ml PBS/I buffer. Keep the flasks on ice.

   The required volume of PBS/I buffer depends on the flask size. For example, use 2 ml for a 25-cm^2 tissue culture flask and 6 ml for a 75-cm^2 tissue culture flask.
3. Add 2 to 6 ml PBS/I buffer to the flask, scrape the cells using a cell scraper, and transfer the suspension to a 50-ml centrifuge tube. Repeat this step with another 2 to 6 ml buffer to ensure complete transfer of the cells.

   Cells grown in suspension are washed in an analogous manner using repetitive centrifugation.
4. Collect the cells by centrifuging 15 min at 2600 × *g* (3000 rpm in Beckman JS-4.2 rotor), 4°C.
5. Discard the supernatant and resuspend the cells in a small volume of PBS/I buffer.
6. Transfer the cell suspension to a labeled cryovial.

   The weight of the empty cryovial can be determined prior to use if the wet weight of the cell pellet is desired as a reference value rather than cell number, radioactivity, or another criterion.
7. Microcentrifuge the cells 15 min at maximum speed, 4°C.
8. Remove and discard the supernatant using a pipettor or Pasteur pipet. Weigh the vial containing the cell pellet. Record the wet weight of the cell pellet (in mg).
9. Freeze the cell pellet in a dry ice/ethanol mixture (optional).

   Frozen cells can be stored at –80°C for at least several months.

Prepare the cell pellets for isoelectrofocusing

10. Retrieve cell pellets from –80°C storage if samples were frozen.
11. Add 400 µl Tris/SDS buffer per 50 to 100 mg cell pellet wet weight. Keep the cells on ice at all times.

    The total amount of protein in the pellet is roughly 5% of the wet pellet weight.
12. Sonicate the sample three times for 3 sec using a sonicator with a microtip at medium power. Keep the samples on ice during sonication.

    Use pulse sonication or wait at least 5 min between sonications. Minimizing heat generation is essential because substantial proteolysis can occur if the sample warms appreciably. If additional sonication is necessary (i.e., if the sample is not homogeneous), let the sample cool down on ice before the next series of sonications.
13. Run a BCA protein assay to determine the protein concentration if sample will be loaded on that basis. For maximum accuracy use the same amount of Tris/SDS buffer in standards as in experimental samples.
14. Prepare labeled cryovials to store aliquots of the sample, if desired. Precool vials on ice prior to making aliquots.

    The amount of protein per aliquot depends on the anticipated future uses of the sample, as repeated freezing and thawing should be avoided. About 500 µg/gel whole-cell extract is

BCU IASI / CENTRAL UNIVERSITY LIBRARY

a maximum load for preparative purposes using 3-mm gels (i.e., isolation of proteins for sequencing or other structural work). Approximately 50 µg/gel is an appropriate load for silver staining. The final protein concentration after completion of the protocol (steps 15 to 20) will equal the concentration found by protein assay divided by 1.1, owing to the addition of reagents after the protein assay step.

15. Add 20 µl DNase and RNase solution per 400 µl Tris/SDS buffer used for sonication (step 11). Incubate 10 min on ice.

16. Add 20 µl of 20% SDS solution and 5 µl of 2-mercaptoethanol per 400 µl Tris/SDS buffer used in step 11. Incubate 5 min at 37°C.

17. Quickly divide samples into previously prepared cryovials and immediately freeze aliquots using a dry ice/ethanol bath. Store at −80°C. Samples stored at −80°C are stable ≥1 year.

    Work quickly to minimize potential proteolysis.

    This step may be omitted if the samples are to be loaded on IEF gels immediately. Generally, the total amount of sample greatly exceeds the amount required for an IEF gel, and freezing aliquots is beneficial. To avoid potential reproducibility problems, all samples should be processed identically.

Prepare the samples for isoelectrofocusing

18. If cell extracts were frozen, thaw samples and immediately add dry urea to 9 M final concentration.

    The amount of urea (in mg) equals 0.83 times the sample volume (in µl). For example, use 83 mg urea per 100 µl sample. The final volume of the sample (with urea added) equals 1.6 times the initial volume (160 µl in the same example).

19. Add an equal volume of lysis buffer (160 µl in the above example) and warm briefly if necessary to dissolve urea.

20. Filter samples using a 0.2-µm microcentrifuge filter unit by microcentrifuging at maximum speed, room temperature, until the entire sample has passed through the filter. Load the desired volume onto the IEF gel.

    If a 500-µg total protein load per 3-mm gel is desired (a practical maximum load for most whole-cell extracts), the protein concentration determined during the protein assay has to be ≥5 µg/µl. If the sample is less concentrated, the sample volume required will be too large for a 3-mm IEF gel. Alternatively, sample loads can be based on cell numbers, radioactivity, or any other appropriate reference (see Basic Protocol 1, step 22).

PREPARING PROTEINS IN TISSUE SAMPLES

SUPPORT PROTOCOL 5

Tissue samples are usually solubilized in lysis buffer using homogenization. After centrifugation, the protein sample can be loaded onto the first-dimension gel. In general, much higher sample-to-sample variability is expected when tissue samples are analyzed.

Materials

Tissue samples
Lysis buffer (see recipe)

Dounce homogenizer or equivalent
Ultracentrifuge and rotor (e.g., Beckman Ti70), 2°C

1. Place tissue sample in a Dounce homogenizer, add ~2 ml lysis buffer per 100 mg tissue, and homogenize the sample on ice (e.g., 3 to 5 strokes).

2. Let the mixture stand a few minutes, then transfer to an appropriately sized ultracentrifuge tube depending on total sample volume.

POLYTECHNIC UNIVERSITY LIBRARY

3. Centrifuge 2 hr at 100,000 × *g* (e.g., 33,000 rpm in a Beckman Ti70 rotor for 100,000 × *g*), or 1 hr at 200,000 × *g* 2°C.
4. Divide the supernatant into aliquots and freeze at −80°C or immediately load an appropriate volume onto the IEF gel.

BASIC PROTOCOL 3

SECOND-DIMENSION ELECTROPHORESIS OF IEF TUBE GELS

Second-dimension gels are identical to those described in *UNIT 10.1* except for sample loading, which requires a broad, flat well. A broad well can be cast using an appropriate two-dimensional comb if the second-dimension gel thickness is slightly larger than that of the first-dimension gel. Alternatively, when the second-dimension gel is being cast, water can be layered over the entire surface of the gel to produce a flat surface that will accommodate the first-dimension gel.

Narrow analytical isoelectrofocusing gels (≤1.5 mm) that fit between the glass plates of the second-dimension gel do not generally require a stacking gel, although a stacking gel may improve resolution under some circumstances. Stacking gels are essential when first-dimension gels >1.5 mm are loaded on reduced-thickness second-dimension gels, for example, when 3-mm first-dimension gels are loaded on 1.5-mm second-dimension gels. To ensure the best reproducibility, casting multiple second-dimension gels in a multigel casting stand is strongly recommended. This is especially important when gradient gels are used for the second-dimension and/or critical comparisons of multiple samples are planned.

This protocol describes all the specific steps required for successfully casting and running the second-dimension gel. The use of beveled plates and an agarose overlay is especially important when 3-mm IEF gels are loaded onto 1.5-mm second-dimension gels.

Materials

2% (w/v) agarose (see recipe)
Equilibration buffer (see recipe)
Isoelectrofocusing gels containing protein samples to be analyzed (see Basic Protocol 1)
Piece of agarose containing molecular weight standards (see Support Protocol 6)

Beveled glass plates
Boiling water bath
Metal or plastic scoop

Additional reagents and equipment for linear and gradient Laemmli gels (*UNIT 10.1*)

Cast the second-dimension gels

1. Assemble the glass-plate sandwich of an electrophoresis apparatus, using a beveled plate for the shorter side of the gel sandwich.

   A beveled plate provides more space for a thicker IEF gel and will accommodate a first-dimension gel that is at least 1 to 2 mm larger than the thickness of the second-dimension gel.

2. If the thickness of the first-dimension gel exceeds that of the second-dimension gel, pour a separating gel of the desired acrylamide concentration and immediately overlay with water to produce a smooth surface. The separating gel height should be a minimum of 2 cm below the top of the beveled plate to accommodate the stacking gel.

BCU IASI CENTRAL UNIVERSITY LIBRARY

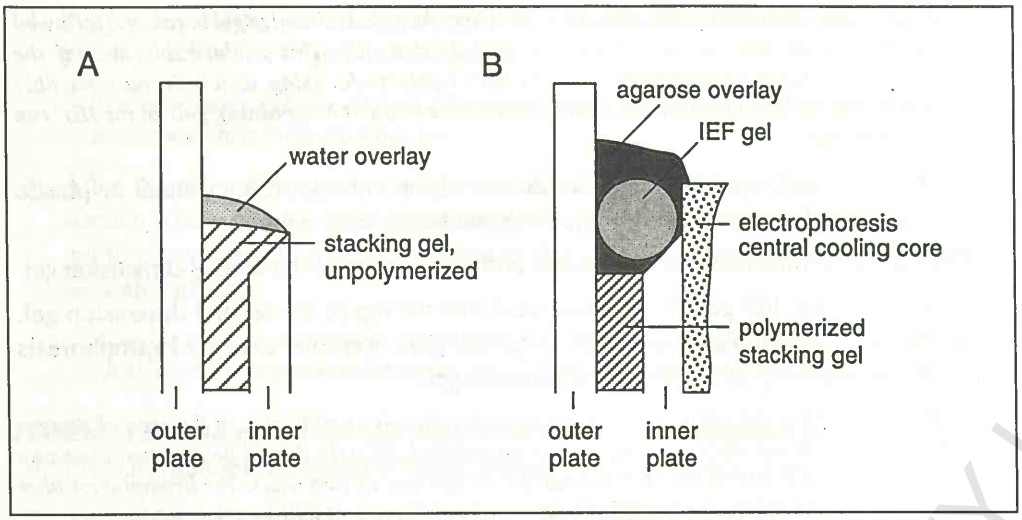


Figure 10.4.2 Casting the second-dimension gel and loading the IEF gel. (**A**) The stacking gel solution should reach to the upper edge of the beveled plate, and then the gel solution has to be overlaid with a minimum volume of water. The water will stay on the surface because of surface tension. (**B**) After polymerization, the gel is mounted on the central cooling core of the electrophoresis unit, and the equilibrated IEF gel is placed on top of the polymerized stacking gel. Excess buffer is removed, and the IEF gel is overlaid with hot agarose/equilibration buffer mixture. After the agarose solidifies, the upper electrophoresis chamber is filled with buffer.

3. After the separating gel has polymerized (a sharp interface between the polymerized gel and the water overlay will reappear), remove the overlay, rinse the gel surface with water, and pour the stacking gel. The stacking gel solution should reach to the top of the bevel. Immediately overlay the stacking gel solution with a minimum amount of water, which will adhere owing to the surface tension (see Fig. 10.4.2A).

   A water overlay of the stacking gel provides a smooth surface and better contact between the IEF gel and second-dimension gel. A small volume of water has to be used to avoid lowering the upper edge of the stacking gel below the edge of the beveled plate. The stacking gel height must be between 1.5 and 2 cm. The solution is filled to the top of the bevel so that after the slight shrinkage that occurs during polymerization the top of the polymerized gel will be near the bottom of the bevel (see Fig. 10.4.2B).

Load the isoelectrofocusing gels onto the second-dimension gels

4. Assemble second-dimension gels in an electrophoresis chamber. Do not pour electrophoresis buffer into the upper chamber.

5. Melt 2% (w/v) agarose in a boiling water bath and add an equal volume of equilibration buffer for use in step 11. Keep agarose/equilibration buffer in the boiling water bath until step 11 is completed.

6. Retrieve isoelectrofocusing gels containing protein samples to be analyzed from storage. Incubate cryotubes containing frozen IEF gels in a 37°C water bath for 15 min for a 3-mm tube gel. A 5- to 7-min incubation is sufficient for 1.5-mm or thinner IEF gels. Do not agitate during thawing, as vigorous agitation of a partially thawed gel can break the gel.

   During this thawing/equilibration step, SDS in the equilibration solution in which the gels were frozen diffuses into the gel matrix and binds to proteins in the IEF gel. The length of incubation in the equilibration buffer is critical because insufficient saturation of proteins with SDS will contribute to vertical streaks on staining. On the other hand, extended incubation in equilibration buffer will result in excessive loss of proteins owing to diffusion of protein out of the gel, which is especially critical for thin IEF gels. For this reason, it is recommended that after extrusion from the IEF tube IEF gels be initially incubated for

BOWMAN-CENTRAL UNIVERSITY LIBRARY

5 min to allow adequate diffusion of glycerol into the gel to minimize gel breakage, followed by freezing on dry ice (see Basic Protocol 1, step 30). This is desirable even if the second-dimension gel will be run directly after isoelectrofocusing, as it is the most feasible way of precisely controlling the equilibration time while the remaining gels in the IEF run are extruded.

7. Pour the gel and equilibration solution out of the cryovial onto a metal or plastic scoop. Carefully remove excess equilibration buffer with a pipet.
8. Place a few milliliters of electrophoresis buffer on the top of the second-dimension gel.
9. Slowly slide the IEF gel off the scoop and onto the top of the second-dimension gel. Remove all air bubbles trapped between the gels. Remove excess electrophoresis buffer from the top of the second-dimension gel.

   The basic end of the gel may be placed on either the left or right side of the second-dimension gel. However, once a convention is established, all gels should be oriented the same way. The acidic end of the IEF gel can be recognized in two ways: the bromphenol blue will usually be yellow, and a bulge (increased gel diameter) will be present.

10. Place a piece of agarose containing molecular weight standards (see Support Protocol 6) beside the basic side of the IEF gel (optional).

    Note that when molecular weight standards are used, the isoelectrofocusing gel has to be shorter than the width of the second-dimension gel.

11. Carefully overlay the IEF gel (and the gel piece with standard proteins) with the hot agarose/equilibration buffer mixture (~2 ml/gel) prepared in step 5. Let the agarose solidify.

    The agarose prevents the IEF gel from shifting position and ensures good contact between the IEF and second-dimension gels.

12. Carefully pour electrophoresis buffer into the upper reservoir, taking care to avoid disturbing the agarose-covered IEF gel.
13. Connect electrodes and run the gels.

    See UNIT 10.1 for electrophoresis conditions.

BASIC PROTOCOL 4

SECOND-DIMENSION ELECTROPHORESIS OF IPG GELS

In this protocol, vertical gel electrophoresis is used as the second dimension for IPG gels in an analogous manner to the protocol described for the second dimension of IEF tube gels (see Basic Protocol 3). One difference is the use of second-dimension gel spacers or gel apparatus that will accommodate an 18-cm-long Immobiline DryStrip; Bio-Rad offers a conversion kit to increase the gel width from 16 cm to 18 cm, and Hoefer Pharmacia offers the Iso-Dalt gel system. The use of beveled plates is not necessary as the 0.5-mm strips are narrower than the second-dimension gel (1.0 or 1.5 mm thick). Another change involves a two-step equilibration of the strips prior to electrophoresis.

Additional Materials (also see Basic Protocol 3)

DryStrip equilibration solutions 1 and 2 (see recipes; prepare fresh in step 4)
Immobiline IPG DryStrip with focused protein (see Basic Protocol 2)
Platform shaker

Cast the second-dimension gel

1. Assemble the glass-plate sandwich of an electrophoresis apparatus, using gel plates wide enough to accommodate an 18-cm-long DryStrip gel.

   Beveled plates are not necessary. If the spacers are not wide enough to accommodate an 18 cm gel, the ends of the gel strip may be trimmed away from the IPG gel so that it will fit on top of the second dimension; however, some very basic or acidic proteins may be lost.

BCU IAS/CENTRAL UNIVERSITY LIBRARY

2. Pour a separating gel of the desired acrylamide concentration and immediately overlay with water to produce a smooth surface.

The separating gel should be a minimum of ~2.5 cm below the top of the inner plate to accommodate a 2-cm stacking gel.

3. After the separating gel has polymerized, remove the water overlay, rinse the gel surface with water to remove any unpolymerized acrylamide, and pour the stacking gel to a height of 0.5 cm from the top of the plate. Overlay with water to produce a smooth surface.

A water overlay provides a smooth surface for better contact between the Immobiline DryStrip and the second dimension gel. The stacking gel height should be ~2 cm.

Load the Immobiline IPG DryStrip gel onto the second-dimension gel

4. Prepare Immobiline DryStrip equilibration solutions 1 and 2 (see recipe).

5. Assemble the second-dimension gels in a electrophoresis chamber. Do not pour electrophoresis buffer into the upper chamber.

6. Melt 2% (w/v) agarose in a boiling water bath. Mix a solution of 1 part 2% agarose to 2 parts equilibration solution 2.

Keep agarose/equilibration buffer mixture in boiling water bath until step 11 is completed.

The agarose prevents the IPG DryStrip from shifting position and ensures good contact between the IEF and second-dimension gels.

7. Using forceps, remove the IPG gels from the electrophoresis tray after isoelectrofocusing is complete or from the –80°C freezer (see Basic Protocol 2, step 19) and place each strip in a separate petri dish with the support film side of the strip facing the petri dish wall. Add 15 ml of DryStrip equilibration buffer 1. Cover and place on a platform shaker for 10 min.

Strips may be run in the second dimension immediately after isofocusing or after storage at –80°C. If the strips have been stored at –80°C, remove them from the freezer, then place in petri dish as stated and continue with the equilibration procedure.

8. Discard equilibration buffer 1 and add 15 ml of equilibration buffer 2. Cover and place on a platform shaker for 10 min.

9. Dampen a piece of filter paper and place on a glass plate. Remove the DryStrips from equilibration buffer 2. Place each strip on its edge on the filter paper to remove any excess buffer.

Strips should not be left in this position for >10 min, or spot sharpness may be affected.

10. Add a small amount of SDS electrode buffer along the glass plate above the second-dimension gel. Place the DryStrip gel in the well with the gel facing out and the basic side to the left. Push the DryStrip down so that it is firmly in contact with the stacking gel of the second-dimension gel. Remove excess running buffer.

11. Overlay the IPG gel strip with the agarose/equilibration buffer (from step 6) and allow agarose to solidify.

12. Carefully pour electrophoresis buffer into the upper reservoir, taking care to avoid disturbing the agarose-embedded IPG DryStrip.

13. Connect electrodes and run the gels.

See UNIT 10.1 for electrophoresis conditions and UNIT 10.5 for gel staining conditions.

BOUVIA SCIENTRA UNIVERSITY LIBRARY

PREPARING MOLECULAR WEIGHT STANDARDS FOR TWO-DIMENSIONAL GELS

Molecular weight markers are usually necessary for the identification of proteins or as references to describe experimental proteins on two-dimensional gels. In many cases, molecular weight markers are required only at the beginning of a project. Once the system is established, common proteins in the sample (e.g., actin or tubulin) provide sufficient references for molecular weight identification on subsequent gels. To minimize any differences in migration of the molecular weight standards and isoelectrofocused proteins, the standard proteins should be loaded on the second-dimension gel in the same manner as the IEF gel. This protocol describes the preparation of standards in solidified agarose. The agarose pieces may be stored at –80°C for at least a year and provide a convenient source of standards for the second-dimension gel. The procedure described is recommended for 3-mm IEF gels. Narrow standards in solidified agarose (made in tubes ≤1.5 mm in diameter) can be prepared by the same method, but extrusion of the thinner agarose gel without breaking is more difficult. The protocol supplies molecular weight markers containing ~2.5 µg of each standard suitable for Coomassie blue staining or 0.25 µg of each standard for silver staining.

Materials

Molecular weight standards (Table 10.1.2)
1× SDS sample buffer (*UNIT 10.1*)
2% (w/v) agarose (see recipe)

Boiling water bath
Glass tubes (3-mm inner diameter)
Plastic or metal tray

1. Prepare 3 ml molecular weight standards in 1× SDS sample buffer using 250 µg of each standard.

   The stated amount is appropriate for Coomassie blue staining of gels. If silver staining is planned, use 25 µg of each standard.

2. Mix the standards with 2 ml of 2% (w/v) agarose melted in a boiling water bath.
3. Prepare clean glass tubes by wrapping one end with Parafilm. Pour the hot mixture into the tubes and let the agarose solidify.
4. Carefully extrude the agarose from the tubes.
5. Cut agarose rods into 5-mm pieces using a razor blade.
6. Freeze all pieces separately on a plastic or metal tray using dry ice.
7. Collect frozen pieces in a plastic bottle and store at –80°C. The standards may be stored ≥1 year.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

DIAGONAL GEL ELECTROPHORESIS (NONREDUCING/ REDUCING GELS)

ALTERNATE PROTOCOL 3

Protein subunit compositions and cross-linked protein complexes can be analyzed by two-dimensional gel electrophoresis using separation under nonreducing conditions in the first dimension followed by reduction of disulfide bonds and separation under reducing conditions in the second dimension. Most proteins will migrate equal distances in both dimensions, forming a diagonal pattern. Proteins containing interchain disulfide bonds will be dissociated into individual subunits and can be resolved in the second-dimension gel.

The approach is similar to that described for two-dimensional gel electrophoresis (see Basic Protocol 3) except, in this protocol, the first-dimension gels are nonreducing (i.e., 2-mercaptoethanol or dithiothreitol is omitted from sample buffer) SDS-denaturing gels instead of isoelectrofocusing gels. Use of 1.2-mm tube gels for the first-dimension separation and 1.5-mm slab gels for the second-dimension run is recommended.

Additional Materials (also see Basic Protocol 3)

Separating and stacking gel solutions (see Table 10.1.1)
1× SDS sample buffer *without* reducing agents (*UNIT 10.1*)
Reducing buffer (see recipe)
1.5% (w/v) agarose in reducing buffer (see recipe; optional, for securing first-dimension gel on second-dimension gel)
Two-dimensional comb (optional)

Additional reagents and equipment for casting tube gels (see Basic Protocol 1), SDS-PAGE (*UNIT 10.1*), and protein staining (*UNIT 10.5*)

Pour and run the first-dimension gel

1. Clean and dry 1.2-mm glass gel tubes for the first-dimension gel (see Basic Protocol 1, step 1).

2. Prepare a separating gel solution with the desired percentage acrylamide (Table 10.1.1); omit the stacking gel for the first dimension.

   Stacking gels can usually be avoided in the first dimension by keeping sample volumes small (i.e., ≤10 µl).

   Less than 200 µl of gel solution is required to cast a single 1.2-mm tube gel 12 cm in length. Adjust the amounts from Table 10.1.1 accordingly.

3. Cast the first-dimension polyacrylamide gels in 1.2-mm tubes using a syringe with a long needle (see Basic Protocol 1, step 5a). Overlay with water and allow the gels to polymerize.

4. Prepare samples in 1× SDS sample buffer without any reducing reagents (i.e., no 2-mercaptoethanol or DTT). Load the samples and electrophorese until the tracking dye is ~1 cm from the bottom of the tube.

Reduce sample and run the second-dimension gel

5. Extrude the gel from the tube (see Basic Protocol 1, steps 27 to 29).

6. Place the extruded gel in a test tube containing 5 ml reducing buffer. Equilibrate 15 min at 37°C with gentle agitation.

7. Cast the second-dimension separating and stacking gels (see Basic Protocol 3, steps 1 to 3), making sure that the top of the stacking gel is at least 5 mm below the top of

BOSTON CENTRAL UNIVERSITY LIBRARY

the short glass plate. Layer water across the entire stacking gel or use a two-dimensional comb.

Most two-dimensional gel combs have a separate small well for a standard or reference sample.

The use of beveled plates (see Basic Protocol 3, steps 1 to 3) is not essential but is still preferred because it will facilitate loading of the first-dimension gel. In this procedure, the first-dimension gel will fit between the glass plates if 1.2-mm tubes are used for the first dimension and 1.5-mm gels are used for the second dimension.

8. Load the first-dimension gel onto the second-dimension gel. Remove any air bubbles trapped between the gels.

   If the first-dimension gel does not remain securely in place, it can be embedded using 1.5% (w/v) agarose in reducing buffer.

9. Carefully pour electrophoresis buffer into the upper electrophoresis chamber and electrophorese using voltages and times appropriate for the gel type selected.

   Parameters for electrophoresis are given in UNIT 10.1.

SUPPORT PROTOCOL 7

USING TWO-DIMENSIONAL PROTEIN DATABASES

Computerized image acquisition and manipulation constitute the only practical method for systematic qualitative and quantitative evaluation of complex protein patterns from different samples that are to be compared by high-resolution two-dimensional gel analysis. Examples of experimental applications include comparisons of tumor cells or tissues with appropriate normal controls and comparisons of a single cell line under different experimental conditions.

There are currently a number of commercially available image acquisition/computer systems specifically designed for comparing two-dimensional gels and storing associated information in a database (*UNIT 10.12*). The systems include both hardware and the necessary software for comparing different gels and producing databases containing the two-dimensional protein patterns, with options for annotating specific spots and producing quantitative comparisons among large numbers of different samples. With most systems, images can be acquired from either stained gels or autoradiographs. The equipment used to obtain two-dimensional gel images includes laser scanners, video cameras, and phosphoimagers. After image acquisition, software running on a microcomputer or workstation is used to refine the image, detect spots, and match spots between different gels.

It is essential that very high-quality, reproducible gels be used for computerized comparisons. The greatest dynamic range in protein abundance for a single two-dimensional gel can be obtained using autoradiography (*UNIT 10.12*) or phosphoimaging. With these methods, up to several thousand spots can be compared and tracked. A representative reference gel or a composite image can be stored and used as a reference for future experiments.

Information related to each spot on the two-dimensional pattern, including the quantity of protein in the indicated spot on different gels used in the comparison, can be archived and updated. Other known information related to a specific spot can also be added to the investigator-built database, including the pI, molecular weight, amino acid composition, sequence, and/or identity of the protein and any other important attributes correlated with the indicated spot. A number of research groups, including those of Garrels and Celis (Garrels, 1989; Garrels and Franza, 1989; Celis et al., 1991), have extensively characterized hundreds of spots from specific cell lines and have used multiple methods to characterize proteins of interest. The most definitive methods for establishing the identi-

BCU IASI / CENTRAL UNIVERSITY LIBRARY

ties for proteins of interest detected by computer-assisted comparisons are protein sequence analysis and, more recently, mass spectrometry of tryptic fragments. Both methods are compatible with the quantities of protein that can be recovered from two-dimensional gels.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

30% acrylamide/0.8% bisacrylamide

30 g acrylamide
0.8 g bisacrylamide
H_2O to 100 ml

Filter solution through 0.2- to 0.45-µm filter (e.g., Micro Filtration Systems, cellulose nitrate, 0.2 µm). Store at 4°C (stable at least 3 months).

Acrylamide is a neurotoxin. Wear gloves and a dust mask when handling solid acrylamide. Wear gloves when working with acrylamide solution. Never pipet acrylamide solutions by mouth.

Agarose in reducing buffer, 1.5% (w/v)

Mix 0.15 g agarose and 10 ml reducing buffer (see recipe). Heat in boiling water bath until dissolved. Prepare immediately before use.

Agarose, 2% (w/v)

Mix 2 g agarose and 100 ml water. Stir on a hot plate until dissolved. Keeping the solution near 100°C, divide by placing 5-ml aliquots in 25-ml glass screw-cap tubes. Let the aliquots solidify. Store at 4°C (stable at least 3 months).

Ammonium persulfate, 2.5% (w/v)

0.25 g ammonium persulfate
H_2O to 10 ml
Prepare immediately before use

DNase and RNase solution

Mix 2.50 ml of 1 M Tris·Cl, pH 7.0 (*APPENDIX 2E*), 250 µl of 1 M $MgCl_2$, (*APPENDIX 2E*), and 2.2 ml water. Add to 5 mg DNase (Worthington) and dissolve DNase. Add 2.5 mg RNase (in solution; Worthington). Mix well, divide into 50- or 100-µl aliquots, and store at –80°C (stable at least 1 year).

The volume of RNase solution needed is variable and depends on the protein concentration, which is reported on the vial label (in mg/ml). The volume of RNase added should be less than 200 µl; if a larger volume is used, the amount of H_2O should be decreased proportionally.

Final concentrations are 0.5 M Tris·Cl, pH 7.0, 0.1 M $MgCl_2$, 0.1% (w/v) DNase, and 0.05% RNase.

DryStrip equilibration solutions

20 ml 1 M Tris·Cl, pH 6.8 (*APPENDIX 2E*)
72 g ultrapure urea
60 ml glycerol
2 g sodium dodecyl sulfate (SDS)
67 ml Milli-Q water
For solution 1: Add 50 mg DTT per 10 ml of equilibration buffer
For solution 2: Add 0.45 g iodoacetamide and a few grains bromphenol blue per 10 ml of equilibration buffer
Make fresh immediately before use

BOSTON / CENTRAL UNIVERSITY LIBRARY

Final concentrations are 50 mM Tris·Cl, pH 6.8; 6 M urea; 30% glycerol; and 1% SDS, in a final volume of 200 ml.

EDTA, 2% (w/v)

2 g Na_2EDTA
H_2O to 100 ml
Adjust to pH 7.0 with NaOH
Store at room temperature (stable several months)

Titrate while dissolving. EDTA is difficult to dissolve without addition of NaOH even when the disodium salt is used.

Equilibration buffer

3 g SDS
7.4 ml 2% (w/v) EDTA, pH 7.0 (see recipe)
10 ml glycerol
2 ml 1.0 M Tris·Cl, pH 8.65 (*APPENDIX 2E*)
0.3 ml bromphenol blue (saturated solution in H_2O)
H_2O to 100 ml
Store at room temperature (stable for several weeks)

Final concentrations are 3% (w/v) SDS, 0.4 mM EDTA, 10% (v/v) glycerol, and 20 mM Tris·Cl, pH 8.65.

Leupeptin, 2 mg/ml

20 mg leupeptin
10 ml water
Divide into convenient volumes
Store at –20°C (stable at least 1 year)

Lysis buffer

2.59 g urea (ultrapure)
1.6 ml H_2O
0.25 ml 2-mercaptoethanol
0.3 ml ampholytes
1.0 ml 20% (w/v) Triton X-100 solution (see recipe)
Prepare immediately before use

Use same ampholytes as for the IEF gel formulation. To dissolve urea, warm the mixture in a 30°C water bath if necessary.

Orthophosphoric acid (H_3PO_4), 0.1 M

13.7 ml 85% phosphoric acid
Water to 2 liters
Make fresh daily

Must be degassed prior to use.

PBS, 10×

152 g NaCl
24 g monobasic sodium phosphate, anhydrous
1600 ml H_2O
Adjust pH to ~6.7 with NaOH
Add H_2O to 2 liters
Store at room temperature (stable at least 1 month)

The 1× solution should be pH 7.3 to 7.5. Final concentrations are 130 mM NaCl and 10 mM sodium phosphate.

BCU IASI CENTRAL UNIVERSITY LIBRARY

PBS with proteolysis inhibitors (PBS/I buffer)

20 ml 10× PBS (see recipe)
20 ml 2% (w/v) EDTA, pH 7.0 (see recipe)
200 µl 0.15 M phenylmethylsulfonyl fluoride (PMSF) in 2-propanol
100 µl 2 mg/ml leupeptin (see recipe)
200 µl 1 mg/ml pepstatin (see recipe)
Adjust to pH 7.2 with HCl
Add H_2O to 200 ml
Prepare immediately before use

Diisopropyl fluorophosphate (DFP) is a better serine protease inhibitor than PMSF at lower temperatures (0° to 4°C); however, although both compounds are toxic, exceptional caution must be exercised with DFP owing to its volatility. If DFP is used, work in a chemical fume hood and carefully follow the supplier's precautions. DFP and PMSF have half-lives on the order of hours in aqueous neutral solutions, and the degradation rate increases rapidly as the pH is increased above neutral. Make aqueous solutions immediately before use. Use of 1 M NaOH is convenient to inactivate residual DFP or PMSF.

Final concentrations are 10 mM sodium phosphate, 130 mM NaCl, 0.2% EDTA, 0.15 mM PMSF; 1µg/ml leupeptin, and 1 µg/ml pepstatin.

Pepstatin, 1 mg/ml

10 mg pepstatin
10 ml anhydrous ethanol
Divide into convenient volumes
Store at –20°C (stable at least 1 year)

Thaw aliquots and mix well immediately before use.

Reducing buffer

0.5 g dithiothreitol (DTT)
0.1 g SDS
1.51 g Tris base
Adjust to pH 6.8 with HCl
Add H_2O to 100 ml
Prepare fresh every time

Final concentrations are 0.5% (w/v) DTT, 0.1% (w/v) SDS, and 125 mM Tris·Cl, pH 6.8.

Tris/SDS buffer

0.3 g SDS
0.6 g Tris base
Adjust to pH 8.0 with HCl
Add H_2O to 100 ml
Divide into 5-ml aliquots
Store at –80°C (stable at least 1 year)

Final concentrations are 0.3% (w/v) SDS and 50 mM Tris·Cl, pH 8.0.

Triton X-100 solution, 20% (w/v)

3 g Triton X-100
12 ml H_2O
Warm in 37°C water bath to dissolve Triton X-100
Store at 4°C (stable ~2 weeks)

Urea, 8 M

0.75 g ultrapure urea
1.0 ml H_2O
Prepare immediately before use

Avoid heating above room temperature.

BOU INST CENTRAL UNIVERSITY LIBRARY

COMMENTARY

Background Information

Two-dimensional gel electrophoresis, using isoelectricfocusing followed by SDS-PAGE, is the single most powerful analytical method currently available for separating complex protein mixtures such as whole-cell or tissue extracts. It is therefore a valuable method for following disease-related changes or for detecting changes in protein expression under diverse experimental conditions. To achieve maximum reproducibility between samples to be compared, multiple gels should be cast and run simultaneously (Anderson and Anderson, 1978a,b).

Despite the exceptionally high resolving power of the method, the total number of proteins that can be resolved in a single two-dimensional gel is only ~1000 to 2000, whereas the total number of proteins in a single mammalian cell type is likely to be at least 10 to 20 times higher. Therefore, the old guideline that a single spot on a two-dimensional gel is a single protein needs to be revised as analytical detection methods improve. Conversely, a single protein (single gene product) can produce multiple, usually adjacent spots in the isoelectrofocusing dimension owing to variable degrees of chemical or post-translational modification. Common examples of variable post-translational modifications that can usually be detected on two-dimensional gels include phosphorylation and glycosylation. Examples of chemical modifications that cause charge heterogeneity include deamidation of side-chain amines (usually asparagines), oxidation of sensitive side chains, and modification of lysines. Potential modification of lysines by urea is particularly important because urea rapidly decomposes to form cyanate, which readily reacts with amino groups, especially above pH 7. Despite potential side reactions, urea is the most useful IEF additive for maintaining solubility of proteins near their isoelectric points.

Combined with electroblotting to PVDF membranes, two-dimensional PAGE has become a valuable preparative tool for protein isolation in addition to its historical role as an analytical method. The sensitivity of many protein analysis methods has improved to the point where one or several spots from two-dimensional gels are sufficient for N-terminal or internal sequence analysis.

Commercially available equipment for running two-dimensional gels can be divided into four groups based on size: microgels (Hoefer Pharmacia Phast system), minigels (e.g., Bio-Rad or Hoefer Pharmacia), standard or full-sized gels (e.g., Bio-Rad or Hoefer Pharmacia), and large or "giant" gels (ESA Investigator 2D gel system or Hoefer Pharmacia Iso-Dalt gel system). In general, the larger the gel, the better the final resolution, but as gel size increases so do costs, difficulty of gel handling, and time requirements.

Standard-size gels provide adequate resolution for most applications and are relatively easy to handle. A 3-mm soluble ampholyte tube IEF gel or a 0.5 mm × 3 mm × 18-cm Immobiline gel has a total protein capacity of ~500 µg for complex protein mixtures such as whole-cell extracts. The maximum capacity for any single protein spot is ~0.5 to 5 µg, depending on the solubility of the protein near its isoelectric point and the separation distance from any near neighbors. A variety of alternative gel sizes, their limits, and their advantages are summarized in Table 10.4.2. The lower protein limit for any of the systems is determined strictly by the available detection methods (*UNIT 10.5*).

Proteins can be detected in two-dimensional gels by the same wide range of techniques used for one-dimensional gels. Autoradiography (*UNIT 10.11*), silver staining (*UNIT 10.5*), and electroblotting to PVDF membranes (*UNIT 10.7*) followed by colloidal gold or colloidal silver staining (*UNIT 10.8*) or immunodetection (*UNIT 10.10*) are among the most sensitive techniques available. If a larger amount of protein is available, Coomassie blue staining of the gel (*UNIT 10.5*) or amido black staining of a PVDF membrane after electrotransfer (*UNIT 10.8*) would be the detection methods of choice.

The major technical limitation in two-dimensional gel electrophoresis is gel-to-gel variation. Even when extreme care is exercised to produce highly reproducible first- and second-dimension gels, some gel-related variability among gels cast at the same time is likely to persist. Another source of variability includes differences in extraction or recovery of proteins during sample solubilization and handling. Maximizing resolution and reproducibility is especially important if computerized comparisons of two-dimensional gels of complex protein mixtures such as cell or tissue extracts are being attempted.

BCU IAS/CENTRAL UNIVERSITY LIBRARY

Table 10.4.2 Size Options for Two-Dimensional Gel Electrophoresis

Gel type	First-dimension gel		Second-dimension gel[a]		Purpose	Comments[b]
	Diameter *D* (mm)	Length *L* (cm)	Thickness *T* (mm)	Height *H* (cm)		
Microgels/minigels[c]	<1.5	<10	<*D*	<10	Analytical	1-4
	<1.5	<10	>*D*	<10	Analytical	3-6
	>1.5	<10	<*D*[d]	<10	Analytical/preparative	1,2,7,8
Full-size gels[e]	<1.5	12-18	<*D*	12-18	Analytical	1-4
	<1.5	12-18	>*D*	12-18	Analytical	3-6
	>1.5	12-18	<*D*[d]	12-18	Analytical/preparative	1,2,7,8
Giant gels[f]	<1.5	>20	>*D*	>20	Analytical	4-6,9
	>1.5	>20	<*D*[d]	>20	Analytical/preparative	1-3,7

[a]The second-dimension gel width has to be at least equal to the IEF tube gel height.

[b]Key to comments: (1) tube gel cannot be placed directly on top of second-dimension gel, and use of agarose is recommended; (2) use of stacking gel is recommended; (3) extrusion and handling are relatively difficult; (4) total protein load is limited to usually ≤50 µg for whole-cell or tissue extracts; (5) tube gel can be placed directly on top of second-dimension gel, and use of agarose is not necessary; (6) use of a stacking gel is not necessary; (7) total protein load capacity is relatively large; (8) extrusion and handling are relatively simple; (9) extrusion and handling are very difficult.

[c]Minigel systems provide rapid separations with moderate resolution. Microgels (Phastgels) are precast gels that are slightly smaller than most minigels.

[d]Use of second-dimension gels thicker than 1.5 mm is generally not recommended owing to difficulty with either efficient staining or efficient electroblotting.

[e]Full-size gels provide resolution satisfactory for most applications.

[f]Giant gels provide very good resolution. Specialized equipment is required, such as Investigator 2D (ESA), Iso-Dalt (Hoeffer Pharmacia), or homemade giant-size gel systems.

Isoelectrofocusing using soluble ampholytes

Soluble ampholytes are mixtures of low-molecular-weight organic compounds with differing side-chain pK_a values that provide buffering capacity. In an IEF gel, the ampholytes migrate to their isoelectric point, where they provide buffering capacity and hence produce stable pH gradients. In theory, any desired pH gradient could be produced by blending ampholytes with appropriate pK_a values. In practice, it is relatively easy to produce pH gradients from ~pH 3.0 or 4.0 to pH 8.0, but stable soluble gradients outside this range are usually not technically feasible. Within these pH limits, some manipulation of the gradient shape and pH range is possible by blending different amounts of specific pH range ampholytes. For example, 0.50 ml of pH 5-7 ampholytes plus 0.25 ml of pH 4-8 ampholytes can be used instead of 0.75 ml of pH 4-8 ampholytes alone to increase the separation distance of proteins in the pH 5.0 to 7.0 range.

Basic Protocol 1 is based on use of 3-mm first-dimension isoelectrofocusing gels and 1.5-mm second-dimension gels using the Bio-Rad two-dimensional gel apparatus (Protean II xi 2D). The method can be easily adapted to equipment from other suppliers or to different-sized gels by adjusting the quantities of reagents used. The protocol uses 8 M urea and Triton X-100 as solubilizing agents. Solubilization of the protein sample applied to the gel as well as maintenance of solubility during electrofocusing are the most critical factors influencing the quality of separation in the first dimension. The most common modification to Triton X-100–based procedures is addition of 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) to the gel and solubilizing buffer mixtures. Addition of SDS to complex samples such as tissue or cell extracts can also enhance reproducible solubilization of the largest possible subset of proteins. Although SDS is charged, in the presence of Triton X-100 it is separated from proteins during focusing and migrates to the acidic end of the gel. Regardless of the method used to maintain solubility, some proteins (especially those >100 kDa) tend to precipitate at pH values approaching their isoelectric point and thus produce horizontal smears on the final second-dimension gel.

BOWMAN GRAY SCHOOL OF MEDICINE LIBRARY

Table 10.4.3 Commercially Available Precast IPG Gels and Immobiline Chemicals[a]

Name	Use	Available pH range
Immobiline DryPlate	Running one-dimensional immobilized pH gradient gels	pH 4.0-7.0, pH 4.2-4.9, pH 4.5-5.4, pH 5.0-6.0, pH 5.6-6.6
Immobiline DryStrip	Running first dimension in two-dimensional gels	
110 mm		pH 4-7L, pH 3-10L
180 mm		pH 4-7L, pH 3-10L[b], pH 3-10NL[c]
Immobiline II	Creating custom gradient immobilized pH gradient gels	pK 3.6, pK 4.6, pK 6.2, pK 7.0, pK 8.5, pK 9.3

[a]From Hoefer Pharmacia.

[b]A linear gradient with maximum resolution above pH 7.0.

[c]A nonlinear gradient with best resolution at pH 5.0-7.0.

Basic Protocol 1 requires several modifications for successful separation of very acidic (Alternate Protocol 1) or very basic (Alternate Protocol 2) proteins. In both cases, prefocusing of the gels must be avoided, and the isoelectrofocusing time has to be reduced. A short separation time does not allow the system to reach equilibrium and is used to establish the desired pH gradient. Separation of very acidic proteins requires modifications of gel and electrode solutions. Separation of very basic proteins has to be performed with the positions of electrodes and electrode solutions reversed in the electrophoresis chamber.

The total protein load per gel depends on the complexity of the sample, the solubility of proteins in the sample, and the diameter of the first-dimension gel. Approximately 4 times as much total sample can be applied to a 3-mm gel compared with a 1.5-mm gel. Another advantage of a larger-diameter isoelectrofocusing gel is that extrusion and gel handling are easier owing to improved strength of the gel. If care is exercised in loading the first-dimension gel onto the 1.5-mm second-dimension gel, the final resolution will be similar to that obtained using smaller-diameter IEF gels. The separating or resolving power of a system is dependent on the quality of ampholytes used, the slope of the pH gradient, and the lengths of both first- and second-dimension gels.

Immobilized pH gradient gels

In immobilized pH gradient (IPG) gels (Basic Protocol 2), the pH gradient is an integral part of the polyacrylamide matrix (Strahler and Hanash, 1991). Because the pH gradient is covalently associated with the polyacrylamide gel matrix, precise, reproducible, and very high-resolution separations can be achieved. A variety of precast gels and all the necessary equipment are commercially available from Hoefer Pharmacia. Reproducible two-dimensional gels can be obtained by running a sample on a narrow strip of immobilized pH gradient gel and then running it in a vertical SDS second-dimension gel. Isoelectrofocusing is performed in a horizontal electrophoresis unit in which multiple gel strips may be run simultaneously. Equipment and reagents are also available for the user to cast custom Immobiline gels in the laboratory.

Some of the major advantages of using precast IPG gels are their ease of use and high reproducibility, the time savings realized, and the fact that precision narrow-range gradients can be used to resolve small charge differences. Because the pH gradient is covalently coupled to the polyacrylamide gel matrix, the pH gradient remains stable and linear during prolonged electrophoresis, thus ensuring reproducibility. This is in contrast to conventional IEF gels, where gradient drift occurs during prolonged electrophoresis. Additionally, the precast gels can be rehydrated in water or in solutions with one or more additives such as urea, CHAPS or Triton X-100, carrier ampholytes, glycerol, and reducing agents, which may help to increase protein solubility. Precast gel strips are available with pH ranges such as pH 3.0 to 10.0 or pH 4.0 to 7.0 as well as narrower ranges (as DryPlates). In addition, a variety of Immobilines permits the user to cast IPG gels with customized pH gradients of any gradient range and shape between pH 3.0 and pH 10.0.

Table 10.4.3 lists types of commercially available gels and Immobilines. With the Immobiline system, the apparent pI of a given protein may be slightly different from that determined by other methods. Therefore, it is

BCU IASE / CENTRAL UNIVERSITY LIBRARY

recommended that a broad-range gradient be tried initially, followed by a narrower-range gradient, if needed.

Diagonal gel electrophoresis

Diagonal gel electrophoresis is a form of two-dimensional analysis useful for investigating the subunit composition of multisubunit proteins containing interchain disulfide bonds (Goverman and Lewis, 1991). Proteins are electrophoresed in the first-dimension in a tube gel (or a slab gel) under nonreducing conditions. The proteins are then reduced in situ, and the first gel (or a strip thereof) is layered onto a second gel and electrophoresed. In the second gel, the proteins migrate at right angles to the original, first-dimension migration. Most cellular proteins are not disulfide-linked and will fall on the "diagonal" in this system; that is, they migrate approximately equal distances in both directions during electrophoresis and lie approximately on the diagonal line connecting opposite corners of the gel. On reduction, component subunits of proteins connected by interchain disulfide bonds will resolve below the diagonal because the individual subunits migrate faster than the disulfide-linked complex during the second electrophoresis. Some proteins with internal disulfide bonds, but no interchain disulfides, may migrate slightly above the diagonal because internal disulfides can produce a more compact molecular shape (causing faster migration in the first dimension).

Investigator 2D gel system

The Investigator 2D gel system was introduced in 1990 by Millipore as the first commercial "large" or "giant" format two-dimensional gel system designed for analytical purposes, although analogous homemade units had been reported earlier (Garrels, 1979; Young et al., 1983). This product line, including precast first- and second-dimension gels, can now be purchased from ESA. The Investigator 2D gel system uses larger gel sizes than the gels described in this unit as well as a number of novel approaches and reagents designed to enhance resolution and gel-to-gel reproducibility. The major features of this system include the following: an increased length of the first-dimension gel (20 cm), an increased length and width of the second-dimension gel (20 × 22 cm), a thread reinforcement of the isoelectrofocusing gel, temperature control during electrophoresis of the second-dimension gel, and use of a special high-tensile-strength acrylamide to facilitate handling of the large, thin second-dimension gel. Narrow gel tubes (1.2 mm) for the first-dimension are standard, but 3-mm-inner-diameter tubes are available to accommodate larger protein loads. The 3-mm IEF gels do not have the thread reinforcement. One modification related to the above protocol includes addition of 0.3% (w/v) CHAPS in the gel solution, which is intended to improve the solubility of proteins and their migration and separation in the gel. The system is supplied with a manual that adequately describes the method.

Critical Parameters and Troubleshooting

Although no individual steps in two-dimensional gel analysis are exceptionally difficult, the large number of steps involved increase the likelihood and possible severity of errors or problems. Several steps are especially critical and may require optimization. The first is sample preparation. Proteins applied to IEF gels have to be completely solubilized. Residual precipitate or even soluble aggregates are likely to cause artifacts on the end of the tube gel or position on the IPG strip where the sample is loaded. Precipitated or aggregated proteins may also interact with soluble proteins, causing components with normally good solubility to coprecipitate or migrate anomalously.

Two general rules of thumb apply to sample solubility: (1) the more complex the sample or the more crude the extract, the more likely problems will be encountered with sample solubility, and (2) the higher the protein load applied to the gel, the more likely solubility problems will arise. Care must also be taken to avoid proteolysis during sample preparation, especially when complex, impure samples are used. If samples are frozen either before or after solubilization, they should be stored at –80°C and should not be subjected to repeated freeze-thawing. Addition of SDS to the solubilization solution increases the solubilization of some proteins; however, SDS should only be used when the IEF gels contain urea and Triton X-100, as the solubilizing agents are required for effective separation of the strongly anionic SDS molecules from the proteins during isoelectrofocusing. Heating of samples in urea-containing solutions must be avoided because urea readily decomposes to cyanate, which reacts with amino groups and causes charge heterogeneity. High concentrations of salts and buffers in the sample should be avoided. Ionic compounds increase the conductivity of the sample and can result in localized overheating, especially for Immobiline gels.

UCSD CENTRAL LIBRARY

Samples analyzed on immobilized pH gradient gels should not contain precipitates. The concentration of salts and buffer ions in the sample should be kept to a minimum (<50 mM) to avoid local overheating of the gel during electrophoresis. If the sample forms aggregates or precipitates at the point of application, apply the sample to a different location (different pH) on the gel.

Early decisions that must be made include the size of the gels required and whether a soluble ampholyte system or an Immobiline gel will be used. The major consideration affecting appropriate gel size is the degree of resolution needed. In general, the smallest gel format should be selected that will provide the needed degree of resolution, because smaller gels are easier, less expensive, and faster to run. Therefore, quick screening of samples or analysis of relatively simple samples can easily be accomplished with microgels or minigels. In contrast, if detailed qualitative or quantitative comparisons of cell or tissue extracts are planned, standard-size or large gels are indicated. Similarly, 3-mm or larger first-dimension tube gels followed by 1.5-mm second-dimension gels in the standard or large format are indicated if the two-dimensional gels will be used for preparative isolation of a protein for applications such as raising antibodies or conducting structural analysis. Immobilized pH gradient gels should very seriously be considered as an alternative to soluble ampholyte gels for most separations owing to their stable and reproducible pH profiles. The IPG gels are particularly appropriate when a narrow pH range is required.

Immobilized pH gradient gels are typically run at 2500 to 3500 V and typically require a focusing time of 16 to 18 hr. Optimal focusing conditions may be experimentally determined by applying the sample to different positions on the gel and estimating the time for the migration patterns to coincide. Some proteins may require longer run times to reach their pI. Because there is no gradient drift, the potential problems with longer run times are limited to sample denaturation or drying out of the gel. These problems usually can be minimized by including a reducing agent in the rehydration solution and/or coating the top surface of the gel with paraffin oil.

The quality of the first-dimensional separation is strongly dependent on the purity of the reagents used, especially the urea and ampholytes. One fairly commonly encountered frustration, when soluble ampholyte gels are used, is that different batches of ampholytes from the same supplier will sometimes produce markedly different two-dimensional gel patterns. Therefore, it is advisable to purchase an adequate supply of a single lot of ampholytes to meet anticipated needs for an entire study where such an approach is feasible. However, whereas ampholytes usually have a reasonably long shelf life at 4°C (usually up to a year), shelf life as well as total ampholyte requirements often cannot be predicted with much certainty.

When any doubt arises about the purity or quality of ampholytes or any other reagent, the reagent should be replaced immediately. Constant monitoring of the system performance, especially when changing lots of ampholytes, urea, or acrylamide, can help minimize potential reagent-associated problems. In most cases, the best standard for a given two-dimensional gel system is an experimental sample or control that is available in sufficient quantity so that many replicate aliquots can be frozen and stored for an extended time at –80°C; alternatively, a sample that can be reproducibly prepared over a long time frame would make an acceptable standard. Such an experimental standard or reference is more likely to detect subtle, but experimentally important, changes in the two-dimensional gel system than commercially available IEF or SDS gel standard mixes.

Another critical factor is equilibration of the first-dimension gel in the second-dimension equilibration buffer. During this step, urea diffuses out of the IEF gel while SDS and reducing reagent diffuse into the gel. If the gel is inadequately saturated with SDS, vertical streaking will result. However, if the gel is incubated in the equilibration buffer for an extended time, a substantial amount of the protein can rapidly diffuse out of the large-pore IEF gel. Losses arising from diffusion can be critical for any experiment because different proteins will diffuse at different rates, but rigorous control of the incubation step is especially important if quantitative comparisons among gels are planned. The simplest method of controlling the incubation time is to freeze the extruded IEF gel after a carefully controlled 5-min incubation in the equilibration buffer; any additional equilibration incubation time required can then be incorporated and carefully controlled when the sample is thawed for loading onto the second-dimension gel.

Another crucial step is loading of the equilibrated IEF gel onto the top of the second-dimension gel. Any irregularity or obstruction between the two gels, including particles of dirt or air bubbles, will affect the flow of current

BCU IAS / CENTRAL UNIVERSITY LIBRARY

and disrupt the resolution of proteins in the final gel. Similarly, poor contact or any movement of the IEF gel during electrophoresis of the proteins out of the IEF gel into the second-dimension gel will lead to artifacts. Therefore, it is advisable to embed the IEF gel in a buffered agarose matrix to ensure good electrical contact between the gels and to prevent gel movement after electrophoresis is initiated.

Finally, the choice of second-dimension gel composition and separation conditions can influence the quality of results. A proper percentage of acrylamide should be selected to optimize resolution within the desired molecular mass range. If gradient gels are needed, use of a multiple gel casting stand is the best way to ensure reproducibility among samples within a single experiment. Further details on optimization of two-dimensional gel systems are presented by Hochstrasser et al. (1988).

When no technical, sample-related, or reagent-related problems are encountered, the final stained two-dimensional gels should contain numerous rounded or slightly elliptical spots. Typically, more than 1000 spots can be detected on a standard 16 × 14–cm gel when using a sensitive staining protocol such as silver staining or autoradiography and a whole-cell extract as a sample. Some horizontal streaks for most high-molecular-mass proteins (proteins exceeding ~100 kDa) are common owing to the decreased solubility of larger proteins near their isoelectric points, even in the presence of urea and nonionic detergent. However, excessive horizontal smearing of proteins smaller than 100 kDa indicates poor isoelectrofocusing, which could be related to one or more of the following factors: sample improperly solubilized or contaminated with interfering substances such as large nucleic acid molecules, poor purity of reagents (check the urea first), poor-quality ampholytes, or insufficient isoelectrofocusing (total volt-hours too low). It is important to note that in general the solubility of any protein is the lowest near its isoelectric point, but there are vast differences among proteins in terms of both the minimum concentration where precipitation becomes a problem and the degree to which precipitation can be prevented by adding different solubilization agents. The best conditions for maintaining solubility during isoelectricfocusing for a given sample type must be determined empirically, although the most universal conditions are those described in the protocols in this unit. In contrast, IEF systems that do not use any detergents or denaturants are limited to that fairly small percentage of proteins which maintain good solubility near their isoelectric point.

If high-molecular-weight proteins are expected but are not present in the final two-dimensional gel, check the sample preparation protocol as well as the sample storage conditions. The most likely problem is proteolysis during sample preparation. Multiple freeze-thawing cycles could contribute to this problem. Vertical smears on the two-dimensional gel suggest (1) insufficient equilibration of the IEF gel (not enough SDS bound to the proteins), (2) poor contact between the IEF and second-dimension gels, or (3) problems related to the stacking gel (too short or wrong buffer). Use of a stacking gel is especially important when large-diameter IEF gels are loaded onto smaller second-dimension gels.

Omission of Triton X-100 or other nonionic or zwitterionic detergent from the final sample loaded on the gel can yield poor results, especially for samples containing SDS, because the amount of detergent in the IEF gels alone may be insufficient to remove bound SDS from proteins. The presence of Triton X-100 in the sample is especially important when SDS sample buffer is used to solubilize protein samples (i.e., after immunoprecipitation). The amount of Triton X-100 in the lysis buffer is normally sufficient for effective dissociation of SDS from proteins. If poor results are encountered with SDS-containing samples, try decreasing the final SDS concentration and/or increasing the final Triton X-100 concentration in the sample.

If no proteins are detected on the gel, check whether (1) the total protein load is appropriate for detection method used, (2) the orientation of electrical connections is wrong or electrical connection during isoelectrofocusing is poor (all gels from that run will be blank), (3) an air bubble obstructs current in a single IEF tube, or (4) the electrical connection is incorrect or is poor during the second-dimension gel separation. Careful monitoring of current and voltage at the beginning, during, and at the end of electrophoretic separations is strongly recommended. Recording the initial and final current and voltage will also facilitate troubleshooting. Additional guidelines for troubleshooting and evaluating artifacts in two-dimensional gel electrophoresis are described by Dunbar (1987).

Anticipated Results

A two-dimensional electrophoretic separation of proteins should produce a pattern of round or elliptical spots separated from one another.

BOSTON CENTRAL UNIVERSITY LIBRARY

The pI range of the separated proteins as well as the observed molecular weight range depend on the first-dimension isoelectrofocusing protocol and the percentage of acrylamide used for the second-dimension gel. A complex protein mixture such as a whole-cell extract should produce more than 1000 silver-stained spots distributed over most of the gel area. Fewer spots will be detected with less sensitive detection techniques such as Coomassie blue staining. On the other hand, separation of radiolabeled proteins and use of multiple exposures permit detection of many low-abundance proteins.

Time Considerations

Time requirements are very dependent on gel size and whether an external cooling unit is used to permit faster separations. Isoelectricfocusing using the standard-size gel format described in Basic Protocol 1 with 3-mm tubes is most conveniently done in an overnight run of ~16 to 18 hr. This separation time can be decreased to ~5 to 6 hr using higher voltages and an external cooling device. Isoelectrofocusing of 18-cm-long IPG gels requires ~16 to 18 hr. Extruding a set of sixteen IEF tube gels and freezing them in equilibration buffer takes 1 to 2 hr, including setup time. Preparing and running SDS gels is described in *UNIT 10.1*. It takes ~30 min to thaw, equilibrate, and load two second-dimension gels.

Overall, if standard-size gels are used without external cooling, it will take ~3 working days before the results of two-dimensional electrophoresis are obtained. A single person can conveniently run about 16 two-dimensional gels in one week, depending on the amount of electrophoresis equipment available. The rate-limiting step in most laboratories is running the second-dimension gels because 16 soluble ampholyte IEF gels or 12 IPG gels can be focused in one run, but loading, running, and detecting results from 12 to 16 second-dimension gels requires substantial operator time and electrophoresis equipment.

Literature Cited

Anderson, N.G. and Anderson, N.L. 1978a. Two-dimensional analysis of serum and tissue proteins: Multiple isoelectrofocusing. *Anal. Biochem.* 85:331-340.

Anderson, N.L. and Anderson, N.G. 1978b. Two-dimensional analysis of serum and tissue proteins: Multiple gradient-slab gel electrophoresis. *Anal. Biochem.* 85:341-354.

Celis, J.E., Rasmussen, H.H., Leffers, H., Madsen, P., Honore, B., Gesser, B., Dejgaard, K., and Vandekerckhove, J. 1991. Human cellular protein patterns and their link to genome DNA sequence data: Usefulness of two-dimensional gel electrophoresis and microsequencing. *FASEB J.* 5:2200-2208.

Dunbar, B. 1987. Troubleshooting and artifacts in two-dimensional polyacrylamide gel electrophoresis. *In* Two-Dimensional Electrophoresis and Immunological Techniques (B.S. Dunbar, ed.) pp. 173-195. Plenum, New York.

Garrels, J.I. 1979. Two-dimensional gel electrophoresis and computer analysis of proteins synthesized by cloned cell lines. *J. Biol. Chem.* 54:7961-7977.

Garrels, J.I. 1989. The QUEST system for quantitative analysis of two-dimensional gels. *J. Biol. Chem.* 264:5269-5282.

Garrels, J.I. and Franza, Jr., B.R. 1989. The REF52 protein database: Methods of database construction and analysis using the QUEST system and characterizations of protein patterns from proliferating and quiescent REF52 cells. *J. Biol. Chem.* 264:5283-5298.

Goverman, J. and Lewis, K. 1991. Separation of disulfide-bonded polypeptides using two-dimensional diagonal gel electrophoresis. *Methods* 3:125-127.

Hochstrasser, D.F., Harrington, M.C., Hochstrasser, A.C., Miller, M.J., and Merril, C.R. 1988. Methods for increasing the resolution of two-dimensional protein electrophoresis. *Anal. Biochem.* 173:424-435.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.

Strahler, J.R. and Hanash, S.M. 1991. Immobilized pH gradients: Analytical and preparative use. *Methods* 3:109-114.

Young, D.A., Voris, B.P., Maytin, E.V., and Colbert, R.A. 1983. Very high resolution two-dimensional electrophoretic separation of proteins on giant gels. *Methods Enzymol.* 91:190-214.

Key Reference

Hochstrasser et al., 1988. See above.

Discusses methods for improving and troubleshooting two-dimensional separation.

Contributed by Sandra Harper,
Jacek Mozdzanowski, and David Speicher
The Wistar Institute
Philadelphia, Pennsylvania

BCU IASI / BIBLIOTECA UNIVERSITARA

Protein Detection in Gels Using Fixation

Detection of proteins in polyacrylamide gels is essential for purification and analysis of proteins. This unit presents basic protocols for detection of proteins by Coomassie blue staining and silver staining. All protocols involve fixing the protein within the gel. Alternate protocols for Basic Protocol 1 detail rapid Coomassie blue staining (Alternate Protocol 1), which is slightly less sensitive, and acid-based Coomassie blue staining (Alternate Protocol 2), which avoids the need for destaining. Basic Protocol 2 for silver staining is accompanied by three alternate protocols: nonammoniacal silver staining (Alternate Protocol 3), which uses more stable solutions; rapid silver staining (Alternate Protocol 4), which involves fixation with formaldehyde instead of glutaraldehyde; and the enhanced-background, two-stage method (Alternate Protocol 5), which employs fixation without cross-linking to enhance sensitivity and also includes an optional reduction step to reduce background staining. A Support Protocol for gel photography is provided as well.

The protocols in this unit and *UNIT 10.6* require proteins that have been previously separated by one- or two-dimensional electrophoresis (*UNITS 10.1-10.4*). Coomassie blue staining is relatively quick and inexpensive, whereas silver staining is considerably more expensive but between 10 and 100 times more sensitive. When subsequent removal of the protein from the gel is desired, detection methods without fixation should be used (*UNIT 10.6*).

COOMASSIE BLUE R-250 STAINING

BASIC PROTOCOL 1

Coomassie brilliant blue R-250 binds nonspecifically to almost all proteins, which allows detection of protein bands in polyacrylamide gels. In this procedure, proteins separated on a polyacrylamide gel are simultaneously fixed and stained in the Coomassie blue staining solution, which turns the entire gel blue. Destaining eliminates the background while the protein bands retain the blue color. Gels can be photographed or dried to maintain a permanent record. Alternatively the "wet" gel can be sealed in a plastic bag with a small amount of destaining solution and stored for many months at 4°C.

Materials

Coomassie blue R-250 staining solution (see recipe)
Polyacrylamide gel containing protein of interest (*UNITS 10.1-10.4*)
Destaining solution: 15% methanol/10% acetic acid (v/v; store up to 1 month at room temperature)
7% (v/v) acetic acid (optional)
Protein-based destaining material (optional): white 100% wool yarn, cellulose sponge, *or* Whatman 3MM filter paper
Gel dryer (optional)

NOTE: Only high-purity water such as Milli-Q-purified water or equivalent should be used throughout the protocol, unless otherwise specified.

1. Add ~10 gel volumes Coomassie blue R-250 staining solution to a plastic or glass container with lid.
2. Remove polyacrylamide gel from the electrophoresis unit and place in staining solution, then place container on a platform shaker and agitate gel slowly to prevent it from adhering to the container.

   CAUTION: *Rapid agitation can result in gel breakage.*

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON GENERAL UNIVERSITY LIBRARY

The appropriate length of time to stain a gel depends on the gel thickness as well as the percentage of acrylamide. In general, gels ≤1 mm thick require only 30 min to 1 hr of staining, whereas gels >1 mm thick require a minimum of 1 to 3 hr of staining. Also, the higher the percentage of acrylamide, the longer it will take to stain the gel. Gels can be left in staining solution overnight without adverse effects.

3. Remove staining solution and rinse the gel briefly with deionized water to remove excess stain.

   Do not reuse staining solution.

4. Add 5 to 10 gel volumes destaining solution and place the container with gel on shaker.

   For more efficient destaining and to reduce waste, a protein-based destaining material such as a piece of wool yarn (6 to 12 in. long), a cellulose sponge (1 in. × 1 in.), or Whatman 3MM filter paper (1 in. × 1 in.) can be added to a minimal amount of destaining solution to bind the dye as it is released from the gel. Alternatively, the destaining solution can be changed several times until the gel shows a clear background with blue protein bands.

   CAUTION: *Overdestaining can occur if the gel is left in destaining solution for a prolonged period (>24 hr) or if the dye-binding material is left with the gel in the destaining solution for more than 2 to 3 hr.*

5. For storage, place gel in a thermally sealed or resealable plastic bag with 1 to 2 ml destaining solution or 7% acetic acid.

   Gels can be stored sealed in plastic bags for several weeks at room temperature and months at 4°C.

6. For a permanent record, photograph the gel (see Support Protocol) or dry the gel as described here. Place a wet piece of Whatman 3MM filter paper on a gel dryer and position the gel (rinsed with water) on top of the filter paper. Cover gel with a piece of plastic wrap and then add a second piece of wet filter paper. Dry 1 to 2 hr at ~80°C or according to gel dryer instructions. As an alternative gel drying method, use a porous membrane to sandwich the gel in a supporting frame, such as BioDesign Gel Wrap. In this case, dry the gel on the countertop at room temperature overnight.

ALTERNATE PROTOCOL 1

RAPID COOMASSIE BLUE G-250 STAINING

This protocol utilizes Coomassie brilliant blue G-250 in place of the Coomassie brilliant blue R-250 used in Basic Protocol 1. Also called xylene brilliant cyanin G, Coomassie brilliant blue G-250 (greenish hue) stains with approximately the same sensitivity as Coomassie brilliant blue R-250 (reddish hue). The G-250 stain produces a somewhat brighter blue coloration than the deep blue of the R-250 stain. The G-250 dye is less soluble than R-250 and is used in a colloidal state rather than as a solution. Due to its lower solubility, G-250 preferentially binds to the protein bands and not to the polyacrylamide gel itself; therefore, little or no destaining is necessary. The procedure requires fixation of gels prior to staining. Protein staining then develops rapidly and can be observed easily through the staining solution. Although the procedure provides faster detection of proteins in gels than Basic Protocol 1, it is often slightly less sensitive, especially with thick gels.

***Additional Materials** (also see Basic Protocol 1)*

Isopropanol fixing solution: 25% isopropanol/10% acetic acid (v/v; store several months at room temperature)
Rapid Coomassie blue G-250 staining solution: 10% (v/v) acetic acid/0.006% (w/v) Coomassie brilliant blue G-250 (Bio-Rad; store several months at room temperature)
10% (v/v) acetic acid

BCU IASTM CENTRAL UNIVERSITY LIBRARY

1. To a plastic container with lid add ~10 gel volumes isopropanol fixing solution. Remove gel from the electrophoresis unit and place in fixing solution. Agitate container on a platform shaker for 10 to 15 min for 1-mm-thick gel or 30 to 60 min for 1.5-mm-thick gel.
2. Discard fixing solution and add 5 to 10 gel volumes rapid Coomassie blue G-250 staining solution. Agitate 30 min to 1 hr for 1-mm-thick gel or 2 to 3 hr for 1.5-mm-thick gel.

   Staining of protein bands becomes apparent within 5 to 15 min.
3. Remove staining solution after appropriate time. Add 5 to 10 gel volumes of 10% acetic acid and agitate slowly. Destain (1 to 2 hr) until a clear background with blue protein bands appears.
4. For storage, place gel in a resealable plastic bag at 4°C with a small amount of 10% acetic acid.
5. To photograph gel, see Support Protocol. To dry gel, see Basic Protocol 1 (step 6).

ACID-BASED COOMASSIE BLUE G-250 STAINING

ALTERNATE PROTOCOL 2

The acid-based Coomassie brilliant blue G-250 staining procedure fixes and stains the gels simultaneously and, unlike the first basic and alternate protocols, does not require a destaining solution. Color intensifies when the gels are placed in water after the staining period, and as little as 0.1 µg protein per band can be detected. Acid-based Coomassie blue G-250 staining has a sensitivity similar to that of the basic Coomassie blue R-250 staining. More heavily loaded protein bands can be detected within 30 min, but ~5 hr is needed to visualize light bands.

Additional Materials (also see Basic Protocol 1)

Acid-based Coomassie blue G-250 staining solution (see recipe)

1. Remove polyacrylamide gel from the electrophoresis unit. If gel contains SDS, incubate in 5 to 10 gel volumes deionized water for a minimum of 5 to 15 min depending on gel thickness to allow SDS to diffuse out.

   SDS will precipitate in the staining solution and produce an opaque background.
2. Place gel into a container with ~100 ml acid-based Coomassie blue G-250 staining solution. Place gel on a platform shaker and agitate gently for 5 hr to overnight.

   Heavily loaded proteins will be visible within 30 min. However, greatest sensitivity is achieved by staining for at least 5 hr.
3. Pour off staining solution and replace with deionized water.

   Even after the gel has been stained overnight, protein bands will not be completely resolved until the gel is placed into deionized water. Bands will appear bright blue against a clear background within 2 to 5 min.
4. For storage, place gel in 1 to 2 ml water in a resealable plastic bag at 4°C.
5. To photograph gel, see Support Protocol. To dry gel, see Basic Protocol 1 (step 6).

BCU/AS/CENTRAL UNIVERSITY LIBRARY

BASIC PROTOCOL 2

SILVER STAINING

Silver staining is about 10 to 100 times more sensitive than Coomassie blue staining; however, results with silver staining vary from protein to protein. After being treated with silver nitrate, the gel is developed until a desired staining intensity is reached. The gel is then fixed with the A component of the Kodak Rapid Fixer reagent. Because the absolute staining intensity is controlled by the length of reaction in the developer and staining intensity is variable for different proteins, estimation of protein amounts using silver staining is generally not advisable unless known quantities of the protein of interest are included on the same gel to calibrate staining intensity.

Materials

Fixing solution A: 50% methanol/10% acetic acid (v/v; store 1 month at room temperature)
Fixing solution B: 5% methanol/7% acetic acid (v/v; store 1 month at room temperature)
10% (w/v) glutaraldehyde
Ammoniacal silver nitrate solution (see recipe)
Citrate developing solution (see recipe)
Kodak Rapid Fixer A

NOTE: Wear gloves that have been thoroughly rinsed with deionized water when handling all materials. Take special care to avoid touching all but the corners of the gel to prevent staining artifacts.

CAUTION: Ammoniacal silver nitrate solution becomes explosive on drying. Discard immediately after use by precipitating the silver chloride with 1 M HCl and washing the mixture down the drain with copious amounts of water, or dispose of solution as required by local regulations.

Fix and stain proteins

1. Add 5 to 10 gel volumes fixing solution A to a plastic or glass container equipped with a lid. After electrophoresis place gel in the container and agitate for 30 min on a platform shaker.
2. Decant fixing solution A, and then add 5 to 10 gel volumes fixing solution B. Fix gel for 60 min with agitation.
3. Decant fixing solution B, and add 5 to 10 gel volumes of 10% glutaraldehyde. Agitate for 30 min in a fume hood.

   CAUTION: *Glutaraldehyde is an irritant. Wear gloves and work only in a fume hood.*
4. Decant glutaraldehyde and wash gel with 5 to 10 gel volumes water for five times, 15 min each time, with agitation. Then add 10 gel volumes water and agitate overnight to remove all traces of glutaraldehyde.
5. Decant water and rinse gel 1 more time with 5 to 10 gel volumes water to remove any residual glutaraldehyde.
6. Decant the last rinse, and add 3 to 5 gel volumes silver nitrate solution. Agitate for 20 min.

Develop stain and process gel

7. Place gel in a clean container and add 5 to 10 gel volumes water. Agitate about 30 sec. Repeat the washing step five times.

BCU IASI CENTRAL UNIVERSITY LIBRARY

8. Add 3 gel volumes diluted citrate developing solution (diluted 1:20 in water). When brown precipitate appears (~30 sec), decant solution, add 5 gel volumes fresh diluted citrate developing solution, and agitate until desired band intensity is attained.
9. After decanting developing solution, add 5 gel volumes diluted Kodak Rapid Fixer A (126 ml in a final volume of 500 ml with water) and agitate for 5 min.

   Kodak Rapid Fixer contains two solutions, Rapid Fixer A and Rapid Fixer B. Only Rapid Fixer A is needed.
10. Decant Rapid Fixer A, add 5 gel volumes water, and agitate for 15 min. Repeat five times.
11. Store gel in a resealable plastic bag at 4°C. Alternatively, photograph gel (see Support Protocol) or dry gel in a gel dryer (see Basic Protocol 1, step 6).

    Silver-stained gels should be photographed as soon as possible because, with time, changes in protein staining and background will be observed.

NONAMMONIACAL SILVER STAINING

ALTERNATE PROTOCOL 3

The nonammoniacal silver staining procedure is similar to Basic Protocol 2 except that more stable solutions are used for staining. Using this procedure, some proteins may be detected that are not recognized by other silver staining methods.

Additional Materials (also see Basic Protocol 2)

5 µg/ml dithiothreitol (DTT; make fresh daily)
0.1% (w/v) silver nitrate
Carbonate developing solution (see recipe)
2.3 M citric acid
0.03% (w/v) sodium carbonate

Fix and stain proteins

1. Remove gel from electrophoresis apparatus and place in a plastic or glass container equipped with a lid. Add 5 to 10 gel volumes fixing solution A. Agitate for 30 min on a platform shaker.
2. Decant fixing solution A, and add 5 to 10 gel volumes fixing solution B. Agitate for 30 min.
3. Remove fixing solution B, add 50 ml of 10% glutaraldehyde, and agitate for 10 min in a fume hood.

   CAUTION: *Glutaraldehyde is an irritant. Wear gloves and work only in a fume hood.*
4. Decant glutaraldehyde solution. Add 5 to 10 gel volumes water and agitate for 15 min. Repeat eight times.

   Alternatively, wash gel overnight in 30 to 40 gel volumes water with slow agitation. The next day add fresh water and agitate for 30 min to ensure that all traces of glutaraldehyde are removed so a clear background is attained.
5. Decant water and add 5 gel volumes of 5 µg/ml DTT. Agitate 30 min.
6. Pour off DTT and, without rinsing, add 5 gel volumes of 0.1% silver nitrate. Agitate 30 min.

BOULDER CENTRAL UNIVERSITY LIBRARY

Develop stain and process gel

7. Decant silver nitrate solution. Rinse gel once for 15 sec with water by filling the container with water and then pouring it out. Then rinse gel quickly two times with a few milliliters carbonate developing solution. Finally, add 100 ml carbonate developing solution and agitate until bands appear as desired.

8. When proper staining has been attained, add 5 ml of 2.3 M citric acid directly to the carbonate developing solution to stop the staining reaction. Agitate for 10 min, replace the solution with water, and agitate for 30 min. Repeat four more times.

   Always use a 1:20 (v/v) ratio of 2.3 M citric acid to carbonate developing solution to neutralize the solution, thus stopping the reaction. If the pH is too high, color development will continue; if the pH is too low, the color will bleach away.

9. To store gel, soak the gel in 5 to 10 gel volumes of 0.03% sodium carbonate for ~1 hr with agitation to prevent bleaching, then place in a resealable plastic bag. Store gel at 4°C.

10. To photograph gel, see Support Protocol. To dry gel, see Basic Protocol 1 (step 6).

ALTERNATE PROTOCOL 4

RAPID SILVER STAINING

Unlike the previous two silver staining procedures (Basic Protocol 2, Alternate Protocol 3) which use glutaraldehyde as a fixative, rapid silver staining employs formaldehyde to fix the gel. The procedure has a sensitivity similar to that of the previous protocols. Also, the developing solution contains thiosulfate to inhibit background staining by dissolving unspecific surface staining. This method is rapid but may not be quite so sensitive as other methods, especially for small proteins.

Additional Materials (also see Basic Protocol 2)

50% methanol/12% acetic acid (v/v)
95% ethanol (optional)
Formaldehyde fixing solution (see recipe)
0.02% (w/v) sodium thiosulfate
0.1% (w/v) silver nitrate
Thiosulfate developing solution (see recipe)
50% methanol

Fix and stain proteins

1. To prevent protein diffusion and improve background, fix gel after electrophoresis, if desired, for a minimum of 1 hr to overnight in 50% methanol/12% acetic acid. Wash gel with 95% ethanol three times, 20 min each time, before proceeding with step 2.

2. Add 3 gel volumes formaldehyde fixing solution. Place container on a platform shaker, then agitate 15 min for a 1-mm-thick gel or 30 min for a 1.5-mm-thick gel.

3. Decant fixing solution, then wash gel by adding 5 to 10 gel volumes water and agitating for 5 min. Repeat two times.

4. Remove water and add 3 gel volumes of 0.02% sodium thiosulfate. Agitate for 1 min.

5. Decant sodium thiosulfate and add 5 to 10 gel volumes water. Agitate for 1 min. Repeat two times.

6. Decant water and add 3 gel volumes of 0.1% silver nitrate. Agitate for 10 min.

BCU IASI CENTRAL UNIVERSITY LIBRARY

Develop stain and process gel

7. Remove silver nitrate, add 5 to 10 gel volumes water, and agitate for 1 min. Repeat two times.
8. Decant water. Add a few milliliters of thiosulfate developing solution until a brown precipitate forms (30 sec), decant, and add 3 gel volumes fresh thiosulfate developing solution.
9. Just before bands reach desired intensity, wash the gel quickly with water to remove excess thiosulfate (to avoid background staining). Stop the reaction by adding 50% methanol/12% acetic acid and agitate for 10 min.
10. Wash the gel with 5 to 10 gel volumes of 50% methanol with agitation for ~30 min to remove acetic acid. Repeat one time.

    The methanol washes are important to prevent darkening of protein bands over a prolonged period.
11. Store gel in a resealable plastic bag at 4°C.
12. For a permanent record, photograph gel (see Support Protocol) or dry gel (see Basic Protocol 1, step 6).

ENHANCED-BACKGROUND (TWO-STAGE) RAPID SILVER STAINING

ALTERNATE PROTOCOL 5

The enhanced-background rapid silver staining method is a two-stage staining procedure that enhances the sensitivity of most proteins and decreases background staining. In this method, however, proteins are not chemically cross-linked by either glutaraldehyde or formaldehyde; hence some proteins, especially those of low molecular weight, may be selectively lost during staining. The first part of the procedure is a rapid silver staining method that is easier and quicker than both basic silver staining and nonammoniacal silver staining. The second part of the procedure, which is optional, improves sensitivity and lessens background staining of the gel.

Additional Materials (also see Basic Protocol 2)

Rapid fixative 1: 40% methanol/10% acetic acid (v/v; store 1 month at room temperature)
Rapid fixative 2: 10% methanol/5% acetic acid (v/v; store 1 month at room temperature)
1× oxidizer (see recipe)
0.17% (w/v) silver nitrate
Thiosulfate developing solution (see recipe)
5% (w/v) acetic acid
Farmers reducer (see recipe)

Rapid Fixation Silver Staining

Fix proteins

1. After electrophoresis, soak gel in 5 gel volumes rapid fixative 1 in a glass or plastic container for 60 min, agitating container on a platform shaker.
2. Pour off used rapid fixative 1 and add 5 gel volumes rapid fixative 2. Agitate for 30 min.
3. Pour off used rapid fixative 2 and add another 5 gel volumes fresh rapid fixative 2. Agitate for 30 min.

BCUMSI CENTRAL UNIVERSITY LIBRARY

Stain and develop gel

4. Pour off used rapid fixative 2 and add 100 ml of 1× oxidizer. Agitate for 10 min.
5. Pour off oxidizer, add 5 gel volumes water, and agitate for 15 min. Repeat four times.
6. Add 100 ml of 0.17% silver nitrate and agitate for 45 min.
7. Wash gel with running water for 2 min.
8. Add 1 gel volume thiosulfate developing solution. After a brown precipitate appears (~30 sec), decant solution and add 5 gel volumes fresh thiosulfate developing solution.
9. After protein bands have reached the desired intensity, stop the reaction by pouring off developing solution and replacing with 5% acetic acid.
10. Pour off acetic acid, add 5 gel volumes water, and agitate 15 min. Repeat three times.
11. To improve background and sensitivity, continue to step 12. Otherwise, store the gel at 4°C in a resealable plastic bag with 5% acetic acid. Alternatively, photograph (see Support Protocol) or dry the gel according to step 6 in Basic Protocol 1.

Improving Sensitivity (Optional)

12. Remove water, add 5 gel volumes Farmers reducer, and agitate for 10 min.

    After this step the gel should be transparent and yellow, and void of any silver-stained bands.

13. Remove reducer and wash gel by adding 5 gel volumes water and agitating for 15 min. Repeat wash five times.

    The gel must be colorless before continuing; therefore, it is preferable to do the last wash overnight.

14. Repeat steps 6 to 10. Store the gel in a sealed plastic bag at 4°C or photograph (see Support Protocol) or dry the gel (see Basic Protocol 1, step 6) for a permanent record.

SUPPORT PROTOCOL

GEL PHOTOGRAPHY

There are several photographic methods that can be used to obtain a permanent record of gels. The choice of method depends on the quality of print desired and access to photographic equipment. The gels must be destained properly for best results; otherwise resolution between background and protein bands will be minimal.

Cameras and film. Fixed-hood (Polaroid DS34) and copy-stand (Polaroid MP-4) cameras, which use instant film, enable fast and easy production of fairly high-quality records. The Polaroid DS34 camera is compact and portable, making it easy to use on the benchtop. The MP-4 camera system is a more complex instrument but is more versatile than the DS34. Both cameras require a light source; for best results a light box that emits even lighting should be used. Use of a deep yellow filter (Kodak Wratten filter no. 8) for Coomassie blue–stained gels and a blue-green filter (Kodak Wratten filter no. 58) for silver-stained gels is necessary for sharp photographic resolution of stained protein bands versus background.

Polaroid type 667 (3¼ × 4¼–in.) black-and-white film is generally used for the DS34 camera, and type 667 or type 57 (4 × 5–in.) black-and-white film is used for the MP-4. For 667 or 57 film an aperture of f16 and a shutter speed of 1/60 to 1/30 sec are recommended; exposure speed is dependent on the light box intensity. If a negative image is desired, Polaroid type 55 (4 × 5–in.) or type 665 (3¼ × 4¼–in.) film can be used with

BCU IAS CENTRAL UNIVERSITY LIBRARY

a shutter speed of ¼ to ½ sec and an aperture of f16. If images on photographs appear too light, the shutter speed or aperture should be increased, and vice versa if the images appear too dark. All settings indicated are approximate, and for best-quality photographs conditions must be determined on a trial basis.

For greatest versatility, 35-mm gel photography can be used with Kodak T-Max 100 or T-Max 400 film. Both films produce negatives with high image quality, but the T-Max 400 film has a higher speed. Photography on 35-mm film also requires a deep yellow filter for Coomassie blue–stained gels and a blue-green filter for silver-stained gels as recommended for the fixed-hood and copy-stand Polaroid cameras.

Contact prints. Kodak X-Omat duplicating film can be used for contact prints as an alternative photography method, especially for silver-stained gels where filters may produce differences in color. However, contact photography does not produce the high-quality images seen with the photographic methods listed above. For contact photography, a piece of glass or plastic is placed between the "wet" gel and the X-ray film in the dark. After the film is briefly exposed using a light source, it is developed in an automated X-ray film processor or manually developed according to the manufacturer's instructions as follows. Develop film with Kodak GBX solution between 20° (5 min) and 26°C (2 min) with no agitation. Then wash with running water for 30 sec at room temperature with continuous agitation. Fix for 2 to 4 min in GBX solution or Kodak Rapid Fixer with intermittent agitation (5 sec every 30 sec), wash with running water for 10 min at room temperature, and let dry.

Gel documentation. Permanent records of stained gels can also be made using a wide range of commercial gel documentation systems. The systems generally involve either a video camera or a densitometric laser scanner that acquires an image which is stored as a digital file in a computer. Computer software allows further processing of the image. In addition to providing digital storage of the image, such systems can produce hard copies of reasonable quality on high-resolution printers.

REAGENTS AND SOLUTIONS

NOTE: Only high-purity water such as Milli-Q-purified water or equivalent should be used to prepare all solutions, unless otherwise specified.

Acid-based Coomassie blue G-250 staining solution

Prepare 0.2% (w/v) Coomassie brilliant blue G-250 (Bio-Rad) in water. Add an equal volume of 2 N H_2SO_4. Mix well and let stand for 3 hr minimum. Filter through Whatman no. 1 filter paper. Measure recovered volume and add 1/9 vol of 10 N KOH. Finally, add 100% (w/v) trichloroacetic acid (TCA) to a final concentration of 12% (w/v). The final concentration of Coomassie blue is ~0.08% (w/v). Store up to several months but maintain solution below pH 1.0.

Ammoniacal silver nitrate solution

Mix 2.5 ml concentrated NH_4OH and 32 ml of 0.36% (w/v) NaOH with stirring, then add H_2O to 150 ml. Continue stirring and slowly add 6 ml of 19.4% (1.2 g/6 ml) silver nitrate for a final concentration of 0.75% (w/v) silver nitrate. A brown precipitate will initially form on adding silver nitrate and will disappear with mixing. Make fresh each time. Silver nitrate stock solution (19.4%) can be stored up to several months at 4°C.

CAUTION: *Ammoniacal silver nitrate solution is explosive on drying. Discard immediately after use by precipitating silver chloride with 1 M HCl, then washing the*

continued

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

mixture down the drain with copious amounts of water (if permitted by local regulations) or disposing of it in another approved manner.

Carbonate developing solution

0.5 ml 37% (v/v) formaldehyde
30 g sodium carbonate (3%, w/v)
H_2O to 1 liter
Make fresh each time

Citrate developing solution

0.5 g sodium citrate
0.5 ml 37% (v/v) formaldehyde
H_2O to 100 ml
Make fresh every time

Coomassie blue R-250 staining solution

Concentrated stock: Mix 24 g Coomassie brilliant blue R-250 (Bio-Rad or Pierce) and 600 ml methanol. Add 120 ml acetic acid. Stir for 2 hr minimum, and preferably overnight. Store up to several months at room temperature.

Staining solution: Mix 1 liter methanol and 60 ml Coomassie blue R-250 concentrated stock. Add 800 ml H_2O, then 200 ml concentrated acetic acid. If precipitation occurs, filter the solution through Whatman no. 1 filter paper prior to use. Store up to several months at room temperature.

The staining solution has approximate final concentrations of 50% (v/v) methanol, 0.1% (w/v) Coomassie brilliant blue R-250, and 10% (v/v) acetic acid.

Farmers reducer

5 g potassium ferricyanide
8 g sodium thiosulfate
H_2O to 500 ml
Make fresh every time

Formaldehyde fixing solution

400 ml methanol
0.5 ml 37% (v/v) formaldehyde
H_2O to 1 liter
Store ≤1 month at room temperature

Oxidizer, 10× and 1×

10× stock:
5 g potassium dichromate
1 ml 70% (w/v) nitric acid
H_2O to 500 ml
Store up to several months at 4°C
1× working solution:
Dilute 10× oxidizer 1/10 with H_2O
Make fresh each time

Thiosulfate developing solution

15 g sodium carbonate
0.5 ml 37% (w/v) formaldehyde
0.2 ml 1% (w/v) sodium thiosulfate
H_2O to 500 ml
Make fresh each time

RCU IASI CENTRAL UNIVERSITY LIBRARY

Background Information

Coomassie blue and silver staining are the two most common approaches used for protein detection in gels with fixation. Silver staining relies on binding of silver to amino acid side chains, primarily sulfhydryl and carboxyl groups, and Coomassie blue binds nonspecifically to proteins. Silver staining is about 10 to 100 times more sensitive than Coomassie blue staining but it is more complex and time-consuming.

The Basic and Alternate Protocols described for Coomassie blue staining differ most significantly in time required. The rapid Coomassie blue G-250 staining procedure (Alternate Protocol 1) uses isopropanol for fixing instead of methanol, and the colloidal stain produces a low background, which allows for faster protein staining (bands appear within 15 min); however, it is not as sensitive as Basic Protocol 1 involving Coomassie blue R-250 (Weber et al., 1972). The acid-based Coomassie blue staining procedure (Blakesley and Boezi, 1977) also uses Coomassie blue G-250, but it does not involve a separate fixation step (Alternate Protocol 2). Acid-based staining achieves a sensitivity similar to that of Coomassie R-250, although for maximum sensitivity gels must be stained 5 hr to overnight.

Basic Protocol 2 for silver staining (Oakley et al., 1980) is a more sensitive protein detection method compared to the nonammoniacal and rapid two-stage alternate silver staining protocols; however, the procedure uses unstable ingredients and tends to require more expensive reagents. Nonammoniacal silver staining (Alternate Protocol 3) is based on a method described by Morrissey (1981). The two-stage silver staining procedure (Alternate Protocol 5) is a modification of the method described by Gorg (1985) and consists of two parts: the first stage, involving rapid silver staining, is followed by an optional series of steps to reduce background. The rapid silver staining procedure (Alternate Protocol 4; Blum et al., 1987) is a quick and easy way to detect proteins within 50 min and, like the alternate two-stage silver staining method, uses thiosulfate to reduce background impurities.

Critical Parameters and Troubleshooting

All solutions used in preparing the polyacrylamide gels as well as the staining solutions must contain high-purity reagents. In addition, gel solutions should be filtered to avoid artifacts in the gel. This is especially important in silver staining, which is based on a photochemical process where impurities can increase background and cause artifactual staining. Powder-free gloves should be used at all times, and care should be taken to touch only the very corners of the gels. One common contaminating protein is skin keratin, which shows up as multiple bands with molecular weights between 55,000 and 60,000. If such bands appear in blank lanes where no protein was loaded, most likely the stacking gel, solubilizing buffer, and/or electrode buffer are contaminated. The problem can be eliminated by avoiding contact of skin and hair with gel and buffer solutions.

Anticipated Results

Silver staining is about 10 to 100 times more sensitive than Coomassie blue staining. The basic silver staining protocol (Basic Protocol 2) enables detection of 1 to 5 ng protein per band, whereas basic Coomassie staining (Basic Protocol 1) has a sensitivity of 0.05 to 0.5 μg protein per band.

Time Considerations

In general, Coomassie blue staining tends to be easier and faster than silver staining. Basic Coomassie staining of standard-sized gels (11 × 16 cm) that are 1.5 mm thick require at least 2 to 3 hr of staining followed by destaining for at least several hours to overnight. Minigels require less time, ~30 min for staining and about the same time for destaining. The rapid Coomassie blue staining procedures (Alternate Protocol 1) allow visualization of protein bands within 5 to 15 min. Silver staining, on the other hand, is very time-consuming because the gels must be treated with many different solutions. Also, gels require successive washes with water to remove glutaraldehyde for good background in the silver staining and nonammoniacal silver staining methods (Basic Protocol 2 and Alternate Protocol 3), and to remove Farmers reducer in the second part of the two-stage silver staining procedure (Alternate Protocol 5). Rapid silver staining (Alternate Protocol 4), however, is the easiest and the most time-efficient of all the silver staining methods as it takes only ~50 min.

BOSTON UNIVERSITY MEDICAL LIBRARY

Literature Cited

Blakesley, R.W. and Boezi, J.A. 1977. A new staining technique for protein in polyacrylamide gels using Coomassie Brilliant Blue G-250. *Anal. Biochem.* 82:580-582.

Blum, H., Beier, H., and Goss H. J. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8:93-99.

Gorg, A. 1985. Horizontal SDS electrophoresis in ultrathin pore gradient gels for the analysis of urinary proteins. *Sci. Tools* 32:5-9.

Morrissey, J.H. 1981. Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. *Anal. Biochem.* 117:307-310.

Oakley, B.R., Kirsch, D.R., and Morris, N.R. 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. *Anal. Biochem.* 105:361-363.

Weber, K., Pringle, J.R., and Osborn, M. 1972. Measurement of molecular weights by electrophoresis on SDS-acrylamide gel. *Methods Enzymol.* 26:3-27.

Key References

Rabilloud, T. 1992. A comparison between low background silver diammine and silver nitrate protein stains. *Electrophoresis* 13:429-439.

Compares many different silver staining methods.

Wilson, C. 1983. Staining of proteins on gels: Comparison of dyes and procedures. *Methods Enzymol.* 91:236-247.

Reviews different Coomassie blue staining procedures.

Contributed by Tara M. DeSilva,
Jeanine A. Ursitti, and David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Protein Detection in Gels Without Fixation

UNIT 10.5 describes protein detection in gels using methods that fix the protein within the gel. Such methods are not always suitable, particularly when protein recovery from the gel is the ultimate goal. Fixatives remove SDS from gels, causing the protein to become entwined in the gel matrix; subsequent extraction from the gel is usually not practical without partial cleavage of the protein or dissolution of the gel matrix. In this unit, methods for the detection of proteins without fixation are described. In general, the sensitivity of the methods is lower than that of protein detection with fixation. Sensitivity is dependent on the volume of the protein band within the gel; hence, sensitivity is the highest for sharp, narrow bands. Techniques described in this unit include basic protocols for protein detection in gels by SDS precipitation, preparation of contact blots, and use of the fluorescent labels IAEDANS and fluorescamine to detect proteins in gels. Several additional methods, including the use of tryptophan fluorescence, guide strips, and minimal protein staining, are discussed in the Commentary.

PROTEIN DETECTION IN GELS BY SDS PRECIPITATION

BASIC PROTOCOL 1

Although proteins can be detected in SDS-containing gels by simply chilling at 0° to 4°C for 3 to 5 hr to precipitate free SDS, resolution is low and special precautions must be taken in preparing and running the gels. The method presented here involves the use of high-ionic-strength solutions to precipitate the SDS and produces the same effect more reliably. Using this method, the SDS in the gel precipitates and forms an opaque background, whereas the SDS bound to the proteins does not precipitate, creating a clear protein band.

Materials

SDS-polyacrylamide gel containing protein of interest (*UNITS 10.1-10.4*)
4 M sodium acetate, pH unadjusted (*APPENDIX 2E*; store at 4°C; stable up to 6 months)
Black, nonreflective surface
Fluorescent desk lamp

NOTE: All steps are performed at room temperature.

1. Immediately after electrophoresis, place the polyacrylamide gel containing the protein of interest into a covered container with 10 to 12 gel volumes of 4 M sodium acetate.
2. Place on a platform shaker at low speed for 40 to 50 min for 1-mm-thick gels or longer for thicker gels.
3. View the gel over a black, nonreflective surface using a fluorescent desk lamp to illuminate the gel from the side. If the gel bands are not yet visible, allow gels to continue to soak.

   Monitor gels about every 15 min, or more often if incubation must continue longer than 1 hr. Be careful not to allow the incubation in high-ionic-strength solution to proceed unattended as the resolution will begin to decrease after a few hours.

4. For recovery of proteins by electroelution, soak gel slices in water to reduce the ionic strength.

   In general, soaking gel slices in 50 ml water for 1 hr, changing the water every 15 min, should be sufficient.

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON CENTRAL UNIVERSITY LIBRARY

BASIC PROTOCOL 2

PREPARATION OF CONTACT BLOTS

Unlike electroblot procedures (*UNIT 10.7*) where it is desired to transfer all of the protein in the gel onto the membrane surface, contact blots are used to create merely a surface replica of the gel by transferring a small amount of protein to the transfer membrane while leaving most of the protein in the gel. The method involves making a gel/membrane sandwich that is soaked in a Tris/glycine transfer buffer for a short time to effect transfer of proteins from the gel to the membrane. The membrane is then stained to visualize proteins. Because the membrane will contain little protein, a high-retention polyvinylidene difluoride (PVDF) membrane should be used, and it must be treated with a sensitive stain such as AuroDye (Amersham) or colloidal gold (*UNIT 10.8*). Protein bands can then be identified in the gel by direct comparison with the transfer membrane.

Materials

Polyacrylamide gel containing protein of interest (*UNITS 10.1-10.4*)
100% methanol
1× transfer buffer (see recipe)

PVDF transfer membrane: 0.2 µm polyvinylidene difluoride (PVDF) membrane (Bio-Rad)
Whatman no. 1 filter paper

Additional reagents and equipment for staining membranes (*UNIT 10.8*)

NOTE: Use powder-free gloves when handling all materials for this procedure, and handle membranes with forceps at the edges to avoid potential artifactual staining.

Prepare transfer membrane and gel

1. Cut PVDF transfer membrane slightly larger than the polyacrylamide gel containing the protein of interest. Cut three slightly larger pieces of Whatman no. 1 filter paper.

2. Wet PVDF membrane in 100% methanol in a clean glass dish for 5 sec. Decant methanol and pour in 1× transfer buffer. Incubate membrane in transfer buffer for at least 5 min.

   The PVDF membrane will wet almost immediately with the 100% methanol. Because the membrane is strongly hydrophobic it will not wet efficiently in either water or transfer buffer alone.

3. Remove polyacrylamide gel from the electrophoresis unit and mark the first lane by cutting off a small piece of the gel on the lower corner of that side. Rinse gel briefly in transfer buffer and place face down into a plastic container with lid.

   The cut edge should now be on the lower right-hand edge of the gel if the gel was loaded from the left-hand side.

Make gel/membrane sandwich

4. Pour ~2 ml transfer buffer on top of the gel.

5. Starting at a place on the gel that has the most buffer on it, lower the prepared PVDF membrane onto the gel. Check carefully for bubbles. Trim the edges of the membrane to the exact size of the gel, including the cut marking the first lane described in step 3.

   Do not cover the gel with too much buffer or the transferred protein bands will appear diffuse on the membrane. Use only enough buffer so that the membrane can be placed on top of the gel without creating bubbles. If a bubble is detected, lift that portion of the membrane to release the bubble and lower the membrane slowly back onto the gel.

ECU IAS / CENTRAL UNIVERSITY LIBRARY

6. Wet the three precut pieces of filter paper with transfer buffer and layer one at a time on top of the membrane. Smooth gently by rolling a test tube or pipet over the surface after adding each layer to eliminate bubbles.

Transfer and stain protein

7. Top with a glass plate to ensure tight contact between the gel and membrane. Cover the plastic container with lid and allow transfer to continue by diffusion for 15 to 30 min.

   The glass plate should be slightly larger than gel/membrane sandwich.

8. After opening the container and removing the glass plate, remove the filter paper from the membrane. To ensure proper realignment after staining the membrane, mark three corners of the gel/membrane sandwich with a syringe needle dipped in permanent ink by piercing through both membrane and gel.

   The membrane and gel should be adequately marked for proper realignment before disassembling the gel/membrane sandwich.

9. Remove the membrane and stain it with a sensitive stain. Keep the gel covered in plastic wrap at 4°C while the membrane is being stained.

10. Once the membrane is stained, identify the position of protein bands in the unstained gel by realigning the gel over the stained blot. Mark bands of interest with a needle dipped in permanent ink as in step 8 or excise the band with a sharp scalpel.

IAEDANS MODIFICATION OF PROTEINS

BASIC PROTOCOL 3

N-Iodoacetyl-*N'*-(5-sulfo-1-naphthyl)ethylenediamine (1,5-I-AEDANS or IAEDANS) is a modified iodoacetic acid (IAA) reagent formed by linkage of IAA to an amino-naphtholsulfonic acid fluorophor. This reagent specifically modifies cysteine residues. IAEDANS-modified cysteines have an excitation maximum of ~340 nm and an emission maximum of 418 nm. In this procedure, modification of protein with the fluorescent label is performed prior to electrophoresis. Several dialysis steps may be needed to provide labeled proteins suitable for electrophoresis. A major advantage of this method is that proteins can be modified in crude mixtures and labeled proteins can be followed through all subsequent steps, including elution from the gel as well as visualization in the gel.

Materials

Protein sample (<5 mg protein in 0.3- to 3-ml volume)
0.75 M Tris·Cl (pH 8.6)/15 mM EDTA (make with 4°C water, measure pH at that temperature, and use within 1 week; store at 4°C)
Urea (Bio-Rad)
Dithiothreitol (DTT or Cleland's reagent; Calbiochem)
Argon or nitrogen gas
N-Iodoacetyl-*N'*-(5-sulfo-1-naphthyl)ethylenediamine (IAEDANS; Sigma; keep yellow crystals desiccated in the dark at –20°C)
2-mercapthoethanol
Additional reagents and equipment for dialysis (*APPENDIX 3B*) and electrophoresis (*UNITS 10.1-10.3*)

CAUTION: Use UV-rated eye protection when working with UV lamps.

BCU ASC CAITRA UNIVERSITY LIBRARY

1. Dialyze 0.3 to 3 ml protein sample (<5 mg total protein) twice against 1 liter Tris/EDTA buffer at 4°C, for a minimum of 3 hr each time.

   The Tris scavenges the cyanate (CN^-) produced by decomposing urea (see below). The indicated volumes are convenient ranges, but the method can be adapted for smaller or larger volumes if needed.

2. Remove protein from the dialysis tubing and measure the volume. Place sample in a tube and add dry urea to a final concentration of 9 M (0.833 g/ml original volume). Bring the solution to room temperature to dissolve urea.

   When the urea is added the sample volume will increase to ~1.5 ml for each 1 ml original volume. It is important to take this into account in calculating amounts of reagents.

3. Immediately add DTT to a final concentration of 20 mM (4.62 mg/1.5 ml reaction solution, based on volume after urea addition).

   DTT will reduce any oxidized components including artificial disulfides that may have formed during protein purification procedures or storage. 2-mercaptoethanol should not be substituted for DTT because it is a weaker reducing reagent.

4. Flush the tube with argon or nitrogen and seal to provide an inert atmosphere. Incubate the tube 1 hr at 37°C.

5. Add 30.1 mg IAEDANS/1.5 ml reaction mixture. Mix to dissolve the reagent, flush with argon or nitrogen, and reseal the tube. Wrap the entire tube carefully in aluminum foil.

   Because IAEDANS is light-sensitive, minimize the time and amount of exposure to light. The indicated quantity of IAEDANS is sufficient to neutralize the DTT (2 moles IAEDANS per mole DTT) plus a moderate excess over thiol groups in the protein. Because IAEDANS modifies thiol groups, this excess is necessary for complete labeling of cysteine residues in the protein.

6. Incubate the sample 1 hr at room temperature. Stop the reaction by adding 2-mercaptoethanol to 100 mM final concentration.

   The sample will appear greenish-yellow in color.

7. Dialyze the protein for ≥4 hr against the desired buffer. Repeat with fresh dialysis buffer until the protein sample is colorless. Conduct the first dialysis in the dark.

   IAEDANS dialyzes away slowly, so each dialysis should be performed for a minimum of 4 hr. Longer times are better. The dialysis can be followed by using a hand-held UV lamp. The protein should be colorless and strongly fluorescent, whereas unreacted reagent is yellow and weakly fluorescent.

   Samples do not have to be completely dialyzed prior to electrophoresis. One dialysis following the labeling is sufficient to reduce the level of unreacted IAEDANS such that it will not interfere with electrophoretic separation; the weakly fluorescent underivatized reagent will migrate near the dye front.

8. Perform gel electrophoresis, following movement of IAEDANS-modified protein during and after the electrophoresis by exposing gels to a hand-held UV lamp. If desired, excise fluorescent bands from the gel with a sharp scalpel and monitor the efficiency of subsequent removal from the gel by exposure to UV light.

BCU IASI CENTRAL UNIVERSITY LIBRARY

FLUORESCAMINE LABELING OF PROTEINS

BASIC PROTOCOL 4

Fluorescamine, or 4-phenylspiro[furan-2(3*H*),1′-phthalan]-3,3′-dione, reacts with α-amino groups of proteins and with the ε-amino groups of lysine residues. The sensitivity of fluorescamine labeling is therefore dependent on the number of lysine residues in a particular protein, but in general as little as 0.5 µg of protein should be easily detected. The excitation maximum for fluorescamine is 390 nm and the emission maximum is 475 nm. This method involves labeling of proteins before they are electrophoresed, like the IAEDANS procedure, but is quicker because only a single round of dialysis is needed, following labeling.

Materials

Protein sample (0.5 to 1 mg/ml)
15 mM sodium phosphate buffer, pH 8.5 (*APPENDIX 2E*; prepare with 4°C water just before use)
SDS (Bio-Rad)
Sucrose
Dithiothreitol (DTT; optional)
1 mg/ml fluorescamine (Sigma) in acetone
5 mg/ml bromphenol blue (optional)
Boiling water bath
Additional reagents and equipment for dialysis (*APPENDIX 3B*) and electrophoresis (*UNIT 10.1*)

CAUTION: Use UV-rated eye protection when using UV lamps.

1. Dialyze 100 µl protein sample against 1 liter of 15 mM sodium phosphate buffer, pH 8.5, at 4°C for a minimum of 3 hr.
2. Remove protein from the dialysis tubing and add dry SDS and sucrose to a final concentration of 5% (w/v) each.

   Alternatively, lyophilize the sample and dissolve the protein (0.5 to 1 mg) in 100 µl of 15 mM sodium phosphate, pH 8.5, containing 5% (w/v) SDS and 5% (w/v) sucrose. Use of Tris should be avoided because it will react with the fluorescamine reagent.

3. Heat the sample 5 min in a boiling water bath. Cool to room temperature.

   If reduction with DTT is necessary to reduce disulfides, add the reducing reagent (20 mM DTT final concentration) prior to heating.

4. Add 5 µl of 1 mg/ml fluorescamine in acetone to the protein mixture and mix well.

   Labeling will occur rapidly, and the degree of modification can be controlled to some extent by the time of incubation prior to electrophoresis. For example, 1 min is sufficient for moderate labeling and 30 min for extensive labeling.

5. Add 5 µl of 5 mg/ml bromphenol blue as a tracking dye, if desired, and apply a 10-µl sample to a polyacrylamide gel well.
6. Conduct electrophoresis. Visualize labeled proteins during or after separation with a hand-held UV lamp.

BCU ASTON CENTRAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use high-purity water such as Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

Transfer buffer

20× stock solution:
200 mM Tris base
2.0 M glycine
Do not adjust pH
Stable up to 1 year at room temperature
1× working solution:
50 ml 20× stock
100 ml methanol
H_2O to 1 liter
Do not adjust pH
Make 1× working solution fresh daily

COMMENTARY

Background Information

Polyacrylamide gel electrophoresis is widely used for the separation of protein mixtures (*UNITS 10.1-10.4*), and a variety of methods to detect the separated proteins have been published. Methods involving fixation of the protein within the gel are presented in *UNIT 10.5*. Fixation is used to inhibit diffusion of the proteins within and out of the gel matrix. Unfortunately, the fixation process eliminates biological activity of the protein. If the protein is needed in its native form and/or recovery of the intact protein from the gel is needed, methods for detection in polyacrylamide gels that do not involve fixation are necessary. In general, the latter methods are required when the protein must be eluted from a polyacrylamide gel. Historically, a standard procedure for locating a desired protein has involved making 1-mm slices of an entire gel, eluting protein from each slice, and assaying for the desired protein. That procedure is time-consuming, and it is likely that the gel will not be cut in the best place. This makes it preferable to visualize specifically the protein in the gel, without fixation. Several methods currently available for this purpose are discussed in this unit.

The choice of label for detecting proteins in polyacrylamide gels without fixation is largely dependent on the ultimate use to which the protein will be put. If it is desirable to modify the protein as little as possible, SDS precipitation and contact blot preparation offer the least manipulation of the proteins (Basic Protocols 1 and 2). The most reliable method for precipitating free SDS uses high-ionic-strength solutions (Higgins and Dahmus, 1979), although chilling gels at 0° to 4°C can be used under certain running conditions (Wallace et al., 1974). Making contact blots involves preparing a gel/membrane sandwich similar to that described for electroblotting (*UNIT 10.7*); however, in a contact blot transfer is achieved by diffusion. The transfer buffer listed in Basic Protocol 2 is approximately half-strength Towbin buffer (Towbin et al., 1979), but other low-ionic-strength transfer buffers may be suitable as well.

Fluorescent labels chemically alter the protein and may affect migration in gels prepared both with and without SDS, but they offer the advantage that the protein can be monitored at all subsequent steps including removal from the gel, which can be problematic. The fluorescent label IAEDANS is especially attractive and can assist in localizing proteins containing cysteine residues. The IAEDANS modification method described in Basic Protocol 3 is based on that of Speicher et al. (1983). Another fluorescent label, fluorescamine, modifies N-terminal amino groups and ε-amino groups on lysine residues; it is preferred over the alternative fluorescent label dansyl chloride because neither its free form nor its hydrolysis products are fluorescent. Labeling conditions are usually selected that minimally modify any individual sites. The simple method of fluorescamine labeling presented in Basic Protocol 4 is based on a procedure described by Eng and Parkes (1974). The fluorescent labels can also be advantageous when the whole protein or peptide products after specific protein cleavage are subjected to HPLC because peaks containing

BCU IASTE TEXTBOOK UNIVERSITY LIBRARY

the desired product can be identified by fluorescence.

In addition to the protocols given in this unit, several alternative simple, nonperturbing methods may be used.

Intrinsic tryptophan fluorescence. In this procedure, the gel is frozen in a suitable container which is floated in liquid nitrogen. Proteins are detected using a hand-held UV lamp. Sensitivity depends on the final temperature achieved, the intensity of the lamp, and the number of tryptophan residues in the protein. In most cases, sensitivity is similar to that obtained with Coomassie blue staining.

Guide strip staining. Multiple identical lanes are run on a polyacrylamide gel. The end lanes of the gel (and the center lane, for improved accuracy) are cut out and stained by one of the methods listed in *UNIT 10.5*. Meanwhile, the unstained portion of the gel is kept in plastic wrap at 4°C. Once the stained portions of the gel have been destained and the size of the gel is again identical (ideally) to that of the unstained portion, the stained gel slices can be lined up with the unstained gel. From this alignment, the region of the desired protein band can be estimated and the corresponding portion of the gel excised.

Minimal protein staining. With this rapid detection method, the gel is treated briefly (1 to 5 min) with Coomassie blue R-250 (*UNIT 10.5*) so that the stain and fixative do not have time to penetrate into the gel matrix, and only the protein molecules near the surface are actually fixed and stained. This method, however, has two drawbacks. The brief staining procedure will emphasize gel surface artifacts, which could lead to incorrect localization of the protein band, and the portion of sample that is fixed in the gel is likely to vary appreciably.

Critical Parameters

In all procedures, it is essential to use high-purity reagents. In addition, if the protein of interest is to be eluted from the gel, it is important to consider all the factors that may affect protein activity after isolation. Although fluorescent probes provide the most sensitive and easiest detection, the modifications must be assessed through the use of appropriate controls to evaluate potential shifts in migration resulting from the chemical modification. When reaction conditions are optimized for complete modification, IAEDANS labeling can be used for two-dimensional electrophoresis using isoelectric focusing (*UNIT 10.4*) with little difficulty (Speicher et al., 1983).

The use of fluorescent probes offers a sensitive alternative for detecting proteins in unstained gels. It is crucial, however, to follow the protocols carefully to obtain efficient and specific labeling of proteins. The modification with IAEDANS is dependent on the pH (pH 8.4 to 8.6 is optimal), state of denaturation (8 to 10 M urea), level of reduction (50- to 200-fold excess of DTT over cysteines in the protein under argon at 37°C for 1 to 2 hr), and the concentration of IAEDANS (allow 2 moles IAEDANS for every mole of DTT plus at least a 10- to 20-fold molar excess of IAEDANS over SH groups). A larger excess of IAEDANS as well as changes in reaction time or pH can lead to modification of other amino acids. Incomplete reduction can lead to incomplete modification. It is therefore important that DTT be used instead of the weaker reducing agent 2-mercaptoethanol. Because modification is performed prior to electrophoresis, incomplete modification of cysteines will result in heterogeneity in isoelectric focusing separations as each modified residue adds a negative charge to the molecule. Molecular weight shifts can also occur, especially when a large number of cysteines are modified in a given protein. A useful alternative approach is to delete the reduction step and/or forgo the use of urea to modify only cysteines not involved in disulfide bonds or solvent-exposed cysteines, respectively. If the reduction step is deleted, the amount of IAEDANS should be reduced to 4.3 mg/1.5 ml of reaction solution (Basic Protocol 3).

The fluorescamine-protein interaction changes the net charge of proteins so that mobilities are changed in polyacrylamide gels run in the absence of SDS. As the charge of the protein is less important in SDS-polyacrylamide gels, moderately labeled proteins should run with similar mobilities to the unlabeled proteins.

Troubleshooting

There are several problems associated with detecting proteins in unstained gels. Perhaps the most common problem arises from the quantity of protein to be detected. With the exception of fluorescent labeling methods to detect proteins, these methods are not very sensitive, and some practice may be required before proteins are detected reliably. One problem that may occur with contact blots is blurring of the transferred bands. If this occurs, it is most likely due to the presence of too much buffer in the transfer sandwich or to insufficient

POLLINS CENTRAL UNIVERSITY LIBRARY

contact between the gel and membrane. Only enough buffer to keep the gel and membrane moist is needed. Improved contact can be achieved by placing a 100-ml beaker filled with water on top of the glass plate. The cover to the gel box must be omitted in this case. Also, if the gel is not rinsed with transfer buffer prior to exposure to the transfer membrane, protein binding to the membrane may be inhibited. On the other hand, if the gel is left in the transfer buffer too long, too much SDS will be removed from the gel and the transfer will not be efficient; therefore, only a brief rinse of the gel is recommended (~10 sec). If insufficient protein is transferred to the membrane by contact blotting, a brief electrotransfer (*UNIT 10.7*) can be used. Transfer time and transfer voltage should be kept as low as consistent with adequate detection sensitivity to ensure that a minimal amount of protein is transferred from the gel.

With the fluorescent labels, if the modification is incomplete or appears to be nonspecific, the reaction conditions should be carefully reviewed. Check calculations for the amount of fluorescent label added as well as reaction pH, reduction conditions, and length of labeling as all these parameters are important for successful labeling of the protein.

Anticipated Results

SDS precipitation and fluorescent labels yield similar sensitivities and should detect as little as 0.5 μg protein. This is only a general estimate, however, as the level of fluorescent labeling is dependent on the number of modified residues in the protein sample. The sensitivity of a contact blot depends on both the amount of protein transferred and the membrane staining technique (see *UNIT 10.8*). The most sensitive staining techniques can detect as little as 2 ng protein.

Time Considerations

The detection of proteins in unstained gels can take as little as 5 min or as long as 1 day, depending on the method. The quickest method available is monitoring of tryptophan fluorescent bands in the gel. SDS precipitation with high ionic strength (Basic Protocol 1) occurs within 10 min for protein bands containing ≥5 μg protein. Contact blots (Basic Protocol 2) are relatively quick to set up and require a short period for protein transfer to occur (30 min). However, the transfer membranes must be stained with a sensitive stain such as AuroDye (Amersham) which takes a minimum of 1 to 4 hr.

With fluorescent labeling, dialysis of the proteins both before and after labeling can consume the majority of time. Dialysis of unlabeled protein into the modification buffer can be done the night before labeling. With fluorescamine (Basic Protocol 4), labeling of the dialyzed protein can be completed within 30 min, and the sample is ready to load on a gel immediately. IAEDANS labeling (Basic Protocol 3) requires more time to complete; although the labeling procedure itself takes only 2 to 3 hr, the reaction mixture should be dialyzed at least 4 hr before the protein is run on a polyacrylamide gel to minimize interference of reaction by-products with the gel separation.

Literature Cited

Eng, P.R. and Parkes, C.O. 1974. SDS electrophoresis of fluorescamine-labeled proteins. *Anal. Biochem.* 59:323-325.

Higgins, R.C. and Dahmus, M.E. 1979. Rapid visualization of protein bands in preparative SDS-polyacrylamide gels. *Anal. Biochem.* 93:257-260.

Speicher, D.W., Davis, G., and Marchesi, V.T. 1983. Structure of human erythrocyte spectrin. II. The sequence of the α-I domain. *J. Biol. Chem.* 258:14938-14947.

Towbin, H., Staehlin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Wallace, R.W., Yu, P.H., Dieckert, J.P., and Dieckert, J.W. 1974. Visualization of protein-SDS complexes in polyacrylamide gels by chilling. *Anal. Biochem.* 61:86-92.

Key Reference

Hames, B.D. and Rickwood, D. (eds.) 1990. Gel Electrophoresis of Proteins: A Practical Approach. Oxford University Press, New York.

A general reference for gel electrophoresis and protein detection methods.

Contributed by Jeanine A. Ursitti,
Tara M. DeSilva, and David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

ECU IAS CENTRAL UNIVERSITY LIBRARY

Electroblotting from Polyacrylamide Gels

Electroblotting of proteins from polyacrylamide gels onto retentive membranes is usually performed to facilitate procedures leading to protein identification and characterization. It has been used for immunoblotting (or Western blotting), N-terminal protein sequencing, in situ protease or chemical cleavage of proteins for further structural analysis (*UNIT 11.2 & UNIT 11.5*), protein band detection using sensitive stains such as AuroDye colloidal gold (*UNIT 10.8*), and protein elution. The electroblotting methods in this unit require that the proteins have already been separated on a polyacrylamide gel (*UNITS 10.1-10.4*). However, if the blotted protein bands will be used for N-terminal protein sequencing or sequencing of cleavage/digestion products, refer to the specific requirements detailed in Alternate Protocol 1, which may affect electrophoresis conditions. Including protein markers or prestained standards in gels is useful because they will transfer to the membrane and can be used to indicate the size of proteins after immunostaining.

This unit contains procedures for electrophoretically transferring proteins onto a variety of membranes including polyvinylidene difluoride (PVDF; Basic Protocol) and nitrocellulose (Alternate Protocol 2) and derivatized membranes. The choice of membrane type for electrotransfer is dependent on the ultimate application for the blot membrane. High-retention PVDF binds proteins tightly and is well suited for applications such as sequencing where high protein amount is critical (Alternate Protocol 1), whereas low-retention membranes may be advantageous for subsequent protein extraction or immunoblot analysis. Most protocols are designed for use with tank transfer setups; Alternate Protocol 3 presents procedures for electroblotting in semidry systems.

In some cases, stained blots are used only to identify protein band patterns while leaving the gel unmodified for subsequent steps (*UNIT 10.8*). If such minimal protein transfer is desired, contact blotting is a suitable alternative (*UNIT 10.6*). This unit also describes procedures for eluting proteins from membranes using detergents (Basic Protocol 2) or acidic extraction with organic solvents (Alternate Protocol 4).

Unless otherwise indicated, the following protocols are primarily designed for electrotransferring proteins from SDS gels (*UNITS 10.1-10.4*). As proteins in other types of gels have much lower charge densities and, sometimes, opposing net charges for different proteins, substantial modifications to the protocols must be made for these specialized applications (*UNIT 10.2*).

ELECTROBLOTTING ONTO PVDF MEMBRANES

BASIC PROTOCOL 1

The following procedure is based on electrotransfer from polyacrylamide gels containing 0.2% sodium dodecyl sulfate (SDS) in the gel solution and electrophoresis and sample buffers. It uses a Bio-Rad solid plate tank transfer apparatus with 7 cm spacing between the electrode plates, which can accommodate gels up to 14 × 18 cm in size and requires ~2.5 liters transfer buffer. The choice of PVDF membrane depends on the planned subsequent use: high-retention Trans-Blot membranes (Bio-Rad) are suitable for protein sequencing, whereas low-retention Immobilon-P membranes (Millipore) are recommended for low-background immunodetection or staining and for recovering proteins from membranes.

BCU MAIN CENTRAL UNIVERSITY LIBRARY

Copyright © 1995 by John Wiley & Sons, Inc.

Materials

1× transfer buffer (see recipe)
Polyacrylamide gel containing proteins of interest (*UNITS 10.1-10.4*)
100% methanol

Powder-free gloves
Electroblotting apparatus: "solid" plate electrode tank transfer system (e.g., Trans-Blot Cell, Bio-Rad)
Glass dishes and trays
Gel support sheet (e.g., porous polyethylene sheet, Curtin Matheson)
PVDF transfer membrane: 0.2 µm PVDF membrane (Bio-Rad) or Immobilon-P (Millipore)
Whatman no. 1 filter paper
Power supply (500 V, 300 mA)

Additional reagents and equipment for staining gels (*UNIT 10.5*)

NOTE: Use powder-free gloves when handling all materials for this procedure and handle membranes with forceps at the edges to avoid potential artifactual staining.

Equilibrate the gel and membrane in transfer buffer

1. Prior to blotting, prepare the blotting apparatus by thoroughly rinsing with high-purity water.

   The optional porous polyethylene sheet is hydrophobic, and it is difficult to get water or buffer into the pores. Be sure to hydrate or "wet" it properly by spraying water under pressure through the sheet or by submerging the entire polyethylene sheet in methanol and then submerging in high-purity water.

2. Fill the transfer tank with 1× transfer buffer. Submerge the gel cassette holder, fiber pads, and polyethylene sheet (after initial hydration as described in step 1) in transfer buffer. Place them inside the tank or submerge in transfer buffer using a separate tray.

   The transfer buffer can be left in a covered transfer tank for up to 2 hr prior to use. Longer times should be avoided because the methanol in the buffer will evaporate, significantly changing the buffer composition.

3. Prepare the polyacrylamide gel containing proteins of interest by disassembling the gel apparatus, removing the stacking gel with a razor blade, and cutting a small piece from the lower left-hand corner of the gel (near lane 1 for gels loaded left to right; see Fig. 10.7.1) to aid in identifying the lanes in subsequent steps.

4. Briefly equilibrate the gel in a glass dish containing ~250 ml of 1× transfer buffer for an appropriate time based on gel thickness (0.5-mm-thick gels for 1 min; 0.75- to 1-mm gels for 5 min; 1.5-mm gels for 15 min).

   Longer equilibrations in transfer buffer prior to electrotransfer may extract too much SDS from the gel and hence reduce transfer efficiency. This equilibration step will remove excess SDS and prevent the gel from swelling during transfer. The negatively charged SDS allows proteins to move toward the transfer membrane in an electric field. On the other hand, too much SDS decreases binding of protein to PVDF membranes. Therefore, the amount of SDS is critical for transfer yield (see Critical Parameters and Troubleshooting).

5. Cut the PVDF transfer membrane about 0.5 to 1 cm larger than the gel. Cut a piece of Whatman no. 1 filter paper ~0.5 cm larger than the PVDF membrane.

6. Wet the PVDF membrane in a clean glass dish containing 100% methanol for 5 sec. Transfer membrane to a tray containing transfer buffer.

BGU IASI / CENTRAL UNIVERSITY LIBRARY

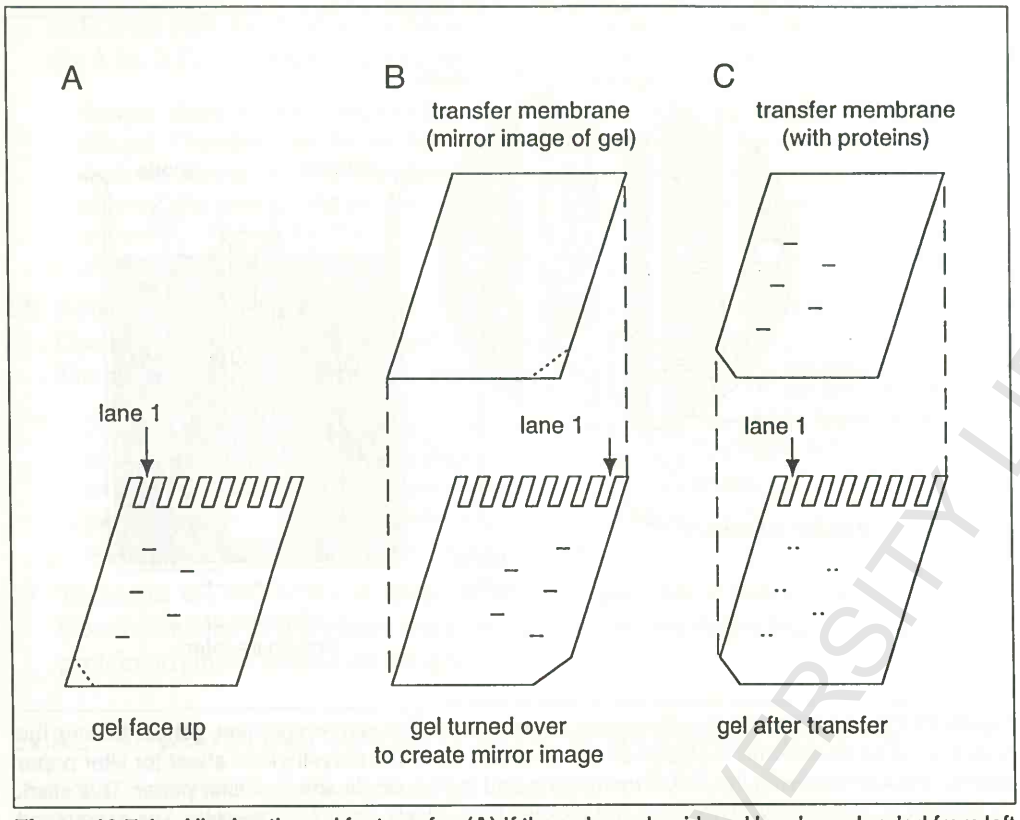


Figure 10.7.1 Aligning the gel for transfer. **(A)** If the polyacrylamide gel has been loaded from left to right, cut a small piece from the corner of the lower left-hand edge of the gel near the first lane. **(B)**When preparing the transfer sandwich, turn the gel over so that the cut edge is on the lower right-hand corner of the gel. This will ensure that the transferred proteins will appear in the same order as in the original gel. **(C)**After transfer, trim the membrane above the cut corner of the gel to mark orientation. Dots indicate where the proteins were before transfer.

The PVDF membrane will wet almost immediately with the 100% methanol. The hydrophobic membrane will not wet in either water or transfer buffer alone. After initial wetting in methanol, do not let the membrane dry; if drying occurs, rewet the membrane in 100% methanol.

Prepare the gel/membrane transfer sandwich

7. Place the opened gel holder cassette on a clean, flat surface. Place one transfer buffer–wetted fiber pad (see step 2) on the top surface, followed by the gel support sheet.

   A sheet of porous polyethylene is preferred as a support because it is rigid and facilitates handling, but filter paper can be used as a simple alternative. If the polyethylene sheet is used, place it with the smooth side up (toward the gel).

8. Place the gel face down on the support, so that the cut edge is now on the right-hand side (see Fig. 10.7.1B). Pour 5 to 10 ml transfer buffer on top of the gel.

9. Remove wet PVDF membrane from the tray containing transfer buffer (step 6). Eliminate air bubbles on both surfaces of membrane by sliding the membrane across the edge of the glass tray. Resubmerge membrane in transfer buffer for a few seconds and then position membrane above gel, letting its center contact the gel, and slowly lower the membrane from the center outward to force any bubbles to the edge of the gel. Rinse gloved hands with high-purity water and smooth the membrane gently to

BOSTON CENTRAL UNIVERSITY LIBRARY

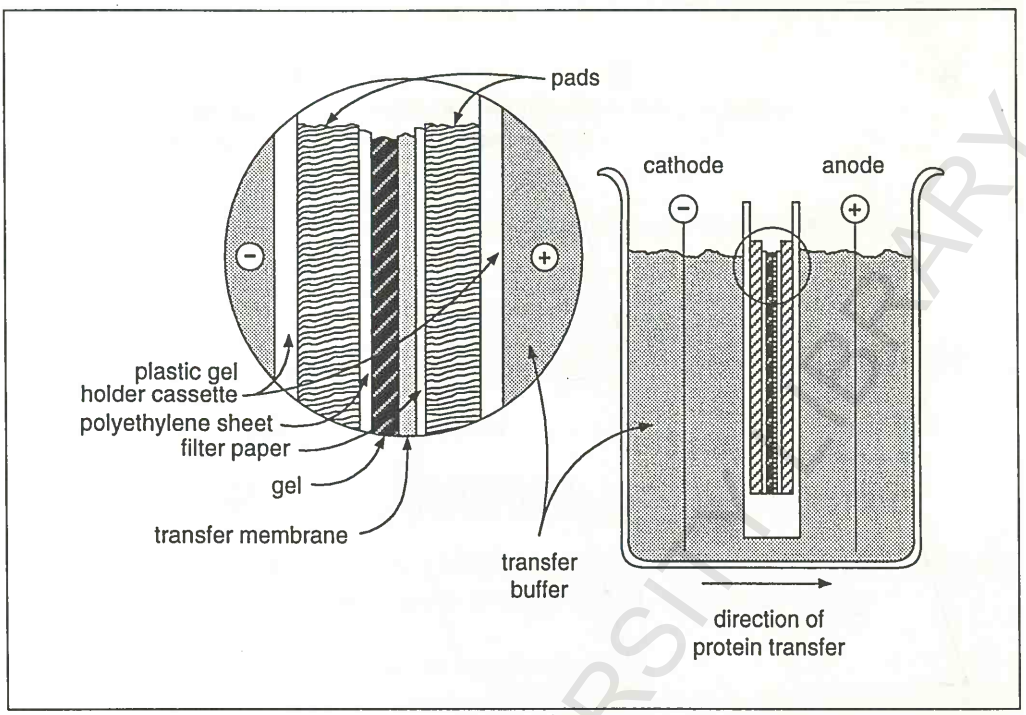


Figure 10.7.2 Electroblotting with a tank transfer unit. The polyacrylamide gel containing the protein(s) to be transferred is placed on the smooth side of the polyethylene sheet (or filter paper sheets) and covered with the PVDF membrane and then a single sheet of filter paper. This stack is sandwiched between two fiber pads and secured in the plastic gel holder cassette. The assembled cassette is then placed in a tank containing transfer buffer. For transfer of negatively charged protein, the membrane is positioned on the anode side of the gel. Charged proteins are transferred electrophoretically from the gel onto the membrane.

ensure uniform contact with the gel. Inspect the membrane for any trapped air bubbles.

This step is extremely important. Air bubbles will prevent the transfer of proteins.

10. Briefly wet the filter paper with transfer buffer and cover the PVDF membrane. Smooth gently to remove any air bubbles. Place the other fiber pad on top of the membrane and close the holder.

Do not allow any area of the PVDF membrane to dry during assembly, or transfer will not occur in these areas. The transfer sandwich should fit together snugly to provide good contact between the membrane and gel. The complete gel sandwich should look like that in Figure 10.7.2.

Conduct protein electrotransfer

11. Slide the assembled transfer cassette into the tank with the gel on the cathode side and the membrane on the anode side.

Proteins in an SDS-polyacrylamide gel have a net negative charge and will migrate toward the anode. For transfer from another type of gel, the orientation of the cassette may need to be reversed.

12. Fill the transfer tank with 1× transfer buffer so that the buffer completely covers the electrode panels but does not touch the electrical connectors.

13. Connect the power supply and perform the transfer at constant current compatible with complete transfer from the gel. For a Bio-Rad Trans-Blot tank transfer apparatus

BCU IASI CENTRAL UNIVERSITY LIBRARY

with solid plate electrodes, use 200 to 250 mA constant current. Transfer 0.5-mm gels for 1 hr, 0.75- to 1-mm gels for 2 hr, and 1.5-mm gels for 3 hr.

Longer times do not appear to increase transfer yield but may lead to overheating of the gel. Transfers can be carried out at constant voltage as well, but the current may increase during the run. Use caution to prevent overheating and/or unacceptably high current. Because of the amount of heat generated at higher transfer rates, it may be necessary to run the transfer in a cold room or with a cooling core to prevent overheating.

14. At the end of the transfer period, turn off the power and disconnect the power supply. Open the transfer sandwich and remove the membrane and gel (Fig. 10.7.1C). Thoroughly rinse the membrane with high-purity water, three times, 5 min each.

The membrane can be stained as described in UNIT 10.8 or used for other detection methods. Stained or unstained membranes can be air dried (for ~30 min) and stored in resealable plastic bags at room temperature for several weeks or at −20°C for a permanent record. Dried membranes can be rehydrated in 50% or 100% methanol and used for subsequent detection procedures.

15. To assess the efficiency of transfer, stain the proteins remaining on the gel with Coomassie blue or with more sensitive silver solutions, depending on the amount of protein originally loaded on the gel.

ELECTROBLOTTING OF PROTEINS FOR SEQUENCE ANALYSIS

ALTERNATE PROTOCOL 1

One-dimensional or two-dimensional gel electrophoresis combined with electroblotting to PVDF membranes provides a method to obtain proteins of high purity that can be used directly for N-terminal protein sequence analysis. The major concern is avoiding possible chemical modification of proteins such as N-terminal blocking. Simple precautions such as casting gels in advance and modifying the composition of reagents to protect proteins minimize the possibility of side reactions during electrophoresis and electroblotting and are described in this protocol.

NOTE: Use high-quality water such as Milli-Q-purified water or equivalent and high-purity electrophoresis reagents (i.e., Bio-Rad) throughout for best results.

Additional Materials (also see Basic Protocol 1)

Thioglycolate (thioglycolic acid, sodium salt; Sigma)
PVDF transfer membrane: 0.2-µm PVDF membrane (Bio-Rad), ProBlott (Perkin-Elmer), or Immobilon-P^{SQ} (Millipore)

Additional reagents and equipment for one- or two-dimensional gel electrophoresis (*UNITS 10.1 & 10.4*) and for staining membranes (*UNIT 10.8*) and gels (*UNIT 10.6*)

1. Cast polyacrylamide gel including stacking gel at least 24 hr but not more than 48 hr prior to use. Store the gel at room temperature until needed, being sure to protect it from dehydration (e.g., store submerged in high-purity water).

For preparing gels choose an acrylamide concentration such that the protein will migrate as a sharp, tight band with an R_f between 0.2 and 0.8. Casting a gel in advance allows complete polymerization, reduces the amount of oxidants and free radicals, and minimizes the possibility of blocking the N terminus or other amino groups of the protein.

2. Add thioglycolate to the electrophoresis buffer in the cathode buffer chamber to a final concentration of 0.1 mM (11.4 mg per liter buffer).

Thioglycolate scavenges remaining free radicals and oxidants in the gel.

BOURNEMOUTH UNIVERSITY LIBRARY

3. Solubilize samples in either 2× or 5× Laemmli solubilizing buffer containing sucrose or glycerol but not urea. Heat to 37°C for 15 min.

   If possible, avoid boiling the sample, as the high temperature increases the risk of chemical modification of the protein. Boiling may be necessary for complete solubilization of some samples, however (e.g., whole viruses).

4. Conduct gel electrophoresis. Proceed with electroblotting (see Basic Protocol 1, steps 1 to 14).

5. After transfer is complete, rinse the membrane six times for 5 min each with a large volume of water (at least 200 ml each time).

   The membrane must be thoroughly rinsed with water immediately after transfer is complete. It is critical that the membrane is not allowed to dry prior to thorough rinsing because even partial drying can prevent effective removal of Tris and glycine.

6. Stain the membrane with amido black or Ponceau S to detect proteins. Stain the gel after transfer as appropriate to judge transfer efficiency.

ALTERNATE PROTOCOL 2

ELECTROBLOTTING ONTO NITROCELLULOSE MEMBRANES

The electroblotting procedure for nitrocellulose membranes differs little from that for PVDF membranes. Nitrocellulose is compatible for use with a moderate amount of SDS (up to 0.1%) in the transfer buffer. The delicate membranes must, however, be handled carefully and protected from high concentrations of organic solvents.

Additional Materials (also see Basic Protocol 1)

0.45-µm nitrocellulose transfer membrane (Schleicher & Schuell)

Additional reagents and equipment for membrane (*UNIT 10.8*) and gel staining (*UNIT 10.6*)

1. Prepare the transfer apparatus and gel as described for PVDF membranes (see Basic Protocol 1, steps 1 to 4).

2. Cut the nitrocellulose transfer membrane to a size slightly larger than the gel and cut a piece of filter paper slightly larger than the membrane.

3. Wet the membrane by slowly introducing the membrane from one corner into a glass dish of 1× transfer buffer. Equilibrate the membrane for 15 min.

   IMPORTANT NOTE: *Never soak the membrane in 100% methanol. During electrotransfer and staining, avoid exposure to high concentrations of organic solvents, which will dissolve the membrane. Up to 20% methanol can be used in the transfer buffer without affecting the membrane.*

4. Proceed with electrotransfer as above for PVDF membranes (see Basic Protocol 1, steps 7 to 13).

5. When transfer is complete, rinse membrane three times for 5 min each with high-purity water; detect proteins on the membrane using an aqueous stain such as Ponceau S if desired.

6. Assess efficiency of transfer by staining the gel after transfer as appropriate.

BCU IASI/CENTRAL UNIVERSITY LIBRARY

PROTEIN ELECTROBLOTTING IN SEMIDRY SYSTEMS

ALTERNATE PROTOCOL 3

An alternative to the tank transfer system is the semidry transfer system. In this procedure, the gel is stacked horizontally on top of the membrane in the transfer apparatus. Because only a small volume of transfer buffer is used, SDS from the gel is less effectively diluted, which may result in incomplete binding and lower yields, especially with PVDF membranes. For this reason, semidry transfer units are not recommended when reproducible high recoveries of electroblotted proteins are desired (e.g., for subsequent sequence analysis). Some procedures recommend stacking multiple transfer sandwiches to achieve several transfers simultaneously. To prevent unbound protein from migrating through the next gel and onto the membrane in the next transfer stack, sheets of porous cellophane sheets or dialysis membrane are placed between adjacent transfer stacks (see Fig. 10.7.3). Semidry electrotransfer requires shorter transfer times than tank transfer.

Additional Materials (also see Basic Protocol 1)

Semidry transfer apparatus
Mylar mask (optional)

Prepare the transfer sandwiches

1. Cut transfer membrane just slightly larger than the gel containing proteins to be transferred, then equilibrate in the appropriate transfer buffer (see Basic Protocol 1, step 6).
2. Cut six sheets of filter paper the same size as the transfer membrane and wet thoroughly in transfer buffer.

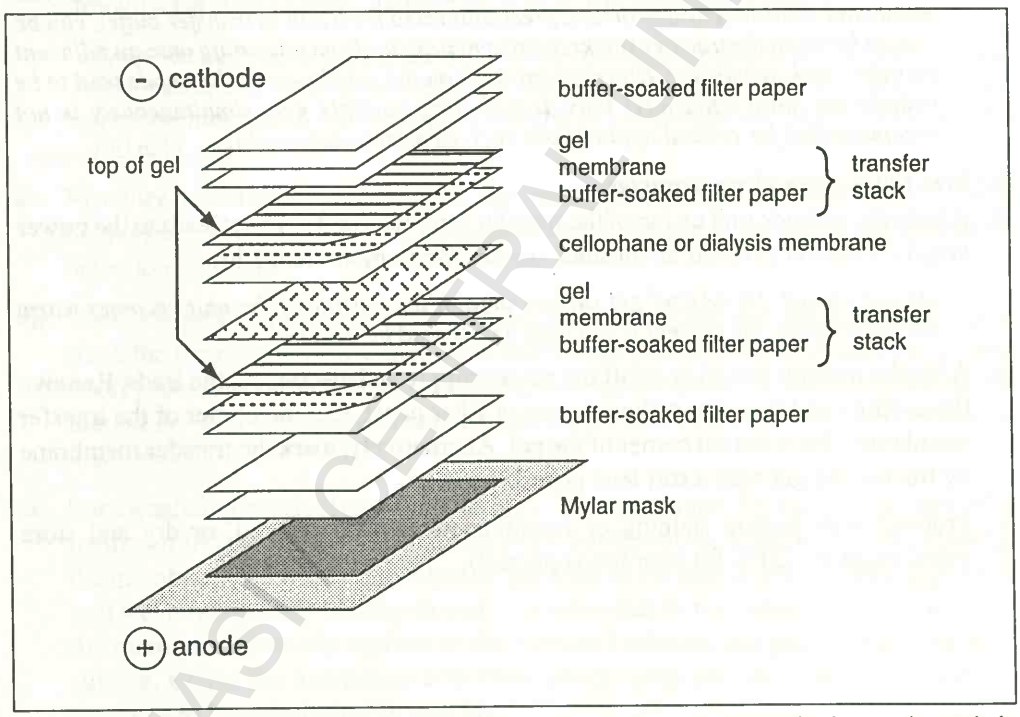


Figure 10.7.3 Electroblotting with a semidry transfer unit. In most cases, the lower electrode is the anode, as shown. Position the Mylar mask (optional) directly over the anode. Layer on three sheets of filter paper that have been wetted in transfer buffer. For negatively charged proteins, place the preequilibrated transfer membrane on top of the filter paper followed by the gel and three additional sheets of wetted filter paper. If multiple gels are to be transferred, separate the transfer sandwiches by inserting a sheet of porous cellophane or dialysis membrane between each stack. Place the cathode on top of the assembled transfer stack(s). Transfer the proteins by applying a maximum current of 0.8 mA/cm^2 gel area.

UCL BIOMEDICAL CENTRAL UNIVERSITY LIBRARY

3. Remove gel containing proteins of interest from the electrophoresis unit. Using a razor blade, cut the lower corner of the gel at the first lane. Equilibrate the gel in transfer buffer (see Basic Protocol 1, step 4).

4. Layer a Mylar mask (optional) followed by three sheets of filter paper on the anode of the transfer unit. Smooth each piece of filter paper to avoid trapping air bubbles between the layers.

   It is important to remove all air bubbles from the transfer stack as they will block the flow of current through that area of the gel, creating blank spots on the transfer membrane.

5. Layer the prepared transfer membrane on top of the filter paper. Gently roll a test tube or Pasteur pipet over the membrane to push out any air bubbles.

6. Place the gel on top of the transfer membrane so that the cut edge is on the left-hand side. With a pair of scissors, cut the lower corner of the membrane even with the gel (to aid in realigning gel and blot after final staining). Again, remove any air bubbles (see Basic Protocol 1, step 9).

   When assembling this transfer sandwich, the gel is placed on top of the membrane, in contrast to the opposite order for tank type sandwich assembly as in Basic Protocol 1. Therefore, it is unnecessary to turn the gel over to obtain the proper orientation (compare Figs. 10.7.1 and 10.7.3).

7. Layer the remaining three sheets of filter paper individually on top of the gel, rolling out any air bubbles after each addition.

   It is possible to transfer multiple gels simultaneously using semidry blotting. As shown in Figure 10.7.3, a sheet of porous cellophane (Hoefer Pharmacia) or dialysis membrane (Bio-Rad or Sartorius), preequilibrated for 5 min in transfer buffer, can be placed between the transfer stacks to prevent proteins from migrating onto an adjacent transfer stack membrane. However, proteins on the gel closest to the anode tend to be transferred more efficiently; thus, transferring multiple gels simultaneously is not recommended for critical applications such as protein sequencing.

Perform the protein electrotransfer

8. Attach the cathode unit on top of the transfer sandwich and connect leads to the power supply. Transfer proteins at constant current for no more than 1 hr.

   Do not exceed 0.8 mA/cm^2 gel surface area. If the outside of the unit becomes warm during transfer, the current is too high and should be lowered.

9. After the transfer period, turn off the power supply and disconnect the leads. Remove the cathode and uppermost three layers of filter paper. Cut the corner of the transfer membrane above the cut corner of the gel. Alternatively, mark the transfer membrane by tracing the gel with a soft lead pencil.

10. Proceed with protein staining or immunodetection, as desired, or dry and store membranes at –20°C for later use (optional).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

PROTEIN ELUTION FROM PVDF MEMBRANES USING DETERGENTS

BASIC PROTOCOL 2

Binding affinities of most proteins to PVDF membranes are relatively strong. The most efficient protocol for protein recovery from PVDF membranes requires the use of detergents, which limits the possible use of extracted samples because detergents are often incompatible with subsequent procedures. This protocol is a simple procedure to elute proteins from the membrane into a Triton/SDS solution. A protocol that employs acidic extraction with organic solvents is also described (Alternate Protocol 4).

Materials

Polyacrylamide gel containing proteins of interest
Ponceau S stain (*UNIT 10.8*; optional)
50% methanol (optional)
Triton/SDS elution buffer (see recipe)

PVDF transfer membrane (Immobilon-P, Millipore)
Transfer apparatus: solid plate electrode tank transfer system (e.g., Trans-Blot, Bio-Rad)

Additional reagents and equipment for staining membranes (*UNIT 10.8*)

1. Prepare gel and membrane for transfer as for PVDF membranes (see Basic Protocol 1, steps 1 to 6).
2. Assemble the gel/membrane transfer sandwich and place in transfer tank (see Basic Protocol 1, steps 7 to 10).
3. Conduct electrotransfer (see Basic Protocol 1, steps 11 to 14), using lower field strengths for step 13 (e.g., transfer for 6 hr at 100 mA for 1.5-mm gels).

   Temperatures >20°C during the transfer can increase the interaction between protein and matrix and make extraction more difficult. If problems arise in eluting proteins from the membrane, try running the transfer in a cold room overnight at low currents (50 mA).

4a. Visualize transferred proteins with any membrane-compatible stain (*UNIT 10.8*). If bound stain will interfere with subsequent steps, use one of the following alternative detection methods.

4b. Visualize proteins with Ponceau S stain (*UNIT 10.8*). After staining with Ponceau S, wash the membrane with water, place under clear plastic wrap, and mark bands with a soft pencil by outlining the protein band through the plastic. Completely destain the membrane with water (a permanent indentation made by the pencil will be left on the membrane).

4c. For visualizing bands after partial drying of the membrane, fill a glass dish with 50% methanol and place the dry membrane on the surface of the solution. Do not submerge the membrane. Identify protein bands (portions of the membrane containing proteins will wet more quickly and dry slower than areas that do not contain protein). Remove the membrane from the surface of the methanol solution and place on a clean glass surface. Cover the membrane with clear plastic wrap and use a pencil to mark the protein bands as described in step 4b.

5. Excise the band(s) of interest by carefully cutting the membrane with a clean razor blade or scalpel.

   Do not allow the membrane to dry during this process. Keep the membrane wet (e.g., submerged in water) at all times (unless using the partial drying procedure).

B.C.U. „M. EMINESCU" IAŞI
CHIMIE

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

6. Place the excised membrane band in a 1.5-ml microcentrifuge tube containing 0.2 to 0.5 ml of Triton/SDS elution buffer for every square centimeter of membrane. Microcentrifuge the sample 10 min at maximum speed, at room temperature, to release the protein from the membrane.

7. Remove supernatant, place it in a new microcentrifuge tube, and microcentrifuge again to remove any particulate material.

Multiple extractions with the elution mixture significantly aid quantitative recovery. In most cases, three extractions remove 90% to 100% of the protein from the membrane.

ALTERNATE PROTOCOL 4

PROTEIN ELUTION FROM PVDF MEMBRANES USING ACIDIC EXTRACTION WITH ORGANIC SOLVENTS

Blotted proteins are usually of high purity and thus good candidates for additional characterization. If the protein cannot be further characterized on the blot and contamination of samples with large amounts of detergent (Support Protocol 1) is unacceptable, acidic elution with organic solvents may be an alternative. The use of organic solvents limits this protocol to PVDF membranes, which have high chemical resistance. This protocol may produce highly variable recoveries for different proteins and is best suited for relatively low-molecular-weight proteins and peptides (<50 kDa).

Additional Materials (also see Basic Protocol 2)

Extraction solution (see recipe)
50% methanol
1:2 (v/v) acetonitrile/formic acid

NOTE: All solutions should be made with high-purity water and can be stored at room temperature for at least 1 month.

1. Rinse the desired number of 1.5-ml microcentrifuge tubes (two tubes per sample) three times with extraction solution, and air dry under an aluminum foil dust cover.

Minimize possible contamination with dust at all times.

2. Visualize and mark the protein bands by partial drying of the membrane using 50% methanol (see Support Protocol 1, step 4c). Use a scalpel or razor blade to cut out the protein bands.

3. Place pieces of membrane containing the protein of interest in a clean microcentrifuge tube.

Combining several bands from multiple gels is usually necessary. Use a pair of tweezers cleaned with 100% methanol to transfer the membrane pieces to the microcentrifuge tube. Avoid contaminating the membrane pieces and microcentrifuge tube with airborne dust or residues from fingers.

4. Add 200 µl extraction solution to each microcentrifuge tube.

An alternative mixture of 1:2 acetonitrile/formic acid seems to be as efficient as the extraction solution described above for some proteins.

5. Agitate tubes for 48 hr at 4°C.

6. Repeat the extraction using 1:2 acetonitrile/formic acid. Combine the extracts.

Multiple extractions (up to three) with the same or different solvent mixtures are often beneficial. The extracts should be combined and kept on ice at all times.

7. Lyophilize protein extracts to provide highly purified samples suitable for most subsequent experiments.

RGU IASL CENTRAL UNIVERSITY LIBRARY

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2E; for suppliers, see SUPPLIERS APPENDIX.

***Extraction solution** (40% acetonitrile/1% trifluoroacetic acid)*

Add 1.0 ml trifluoroacetic acid to 59 ml high-purity water, then add 40 ml acetonitrile. Store at room temperature. Solution is stable for at least 1 month in a tightly sealed container.

Transfer buffer, 20×

200 mM Tris base
2.0 M glycine
Do not adjust pH of final solution (stable at room temperature for at least 1 month)

(Prepare 1× transfer buffer by mixing 200 ml 20× stock, 400 ml methanol, and water to 4 liters.) Prepare 1× transfer buffer on the same day of use. Keep at room temperature.

Triton/SDS elution buffer

50 mM Tris·Cl, pH 9.0 (*APPENDIX 2E*)
2% (w/v) SDS (*APPENDIX 2E*)
1% (v/v) Triton X-100

Prepare solution on the same day of use. Keep at room temperature.

COMMENTARY

Background Information

The transfer of proteins from polyacrylamide gels onto blot membranes offers many benefits. Transferred proteins can be eluted from the membrane, probed with antibodies (immunoblotting), used for N-terminal protein sequencing as well as other structural analyses including amino acid analysis, protease cleavage (*UNIT 11.2*), and chemical cleavages (*UNIT 11.5*), or stained by a variety of highly sensitive techniques. In addition, several manipulations can be performed on the same blot by loading samples in multiple lanes on the gel before electroblotting.

Prestained protein standards are useful guides for cutting the membrane prior to different analysis procedures, such as using one part for immunoblotting and the other for protein staining. Prestained radiolabeled standards are useful for aligning the blot membrane with autoradiograms. Prestained standards will migrate differently on the gel compared with the unmodified standard proteins, usually showing higher molecular weight.

During electrotransfer, proteins migrate out of gels in an electric field according to the charge on the protein. Most electrotransfers employ a tank transfer apparatus (Fig. 10.7.2) in which the gel/membrane transfer sandwich is mounted in a cassette and placed in a tank of Tris/glycine/methanol transfer buffer. The semidry transfer system (Fig. 10.7.3) is an alternative. The major advantages of the semidry system are its use of substantially less buffer and the shortening of transfer time due to the close positioning of the electrodes, which produces a high field strength with minimal heating. A variety of transfer membranes, which can effectively bind proteins based on different types of protein-membrane interactions, are commercially available. These include PVDF, nitrocellulose, and nylon membranes as well as a number of derivatized membranes.

PVDF membranes bind proteins primarily through hydrophobic interactions and are commonly used for their chemical resistance as well as physical stability. High-affinity PVDF membranes such as Trans-Blot (Bio-Rad), Pro-Blott (Perkin-Elmer), and Immobilon-P^{SQ} (Millipore) are preferred for blots intended for use in N-terminal protein sequencing, whereas low-retention membranes such as Immobilon-P (Millipore) may produce lower background in both immunoblotting and common staining procedures. In addition, low-retention membranes are preferred when proteins will be extracted from the membrane.

Nitrocellulose binds proteins primarily by hydrophilic and/or electrostatic interactions and for this reason is less sensitive than PVDF

BYU SOUTHERN UNIVERSITY LIBRARY

membranes to the concentration of SDS in both the gel and transfer buffer. Its tolerance to SDS allows relatively good recovery of poorly migrating proteins because higher concentrations of SDS can be left in the gel or added to the transfer buffer without adversely affecting protein binding to the membrane. In addition, nitrocellulose is often used for immunoblots, although PVDF membranes have replaced nitrocellulose for this application in many laboratories. The major limitations of nitrocellulose are its poor binding of low-molecular-weight proteins and peptides, mechanical weakness, and lack of resistance to organic solvents.

PVDF can be derivatized to modify the nature of protein-membrane interactions and to potentially increase the strength of protein binding. Derivatization most commonly involves adding a charge to the membrane so that proteins will be held on the membrane by electrostatic charges as well the usual hydrophobic or hydrophilic interactions. The procedure for electrotransfer onto derivatized membranes is essentially identical to that for PVDF or nitrocellulose membranes described in this unit. Although derivatized membranes are better for binding RNA and DNA, most derivatized membranes have not offered clear advantages for blotting proteins.

One commercially available derivatized PVDF membrane that has potentially useful binding properties for proteins is Immobilon-CD (Millipore), which has been derivatized to produce a cationic surface so that it behaves like an ion-exchange support (Patterson et al., 1992). In addition, the membrane is treated to prevent nonspecific protein adsorption. The well-defined nature of the protein/Immobilon-CD interaction allows good recoveries of electroblotted proteins under relatively mild conditions. Proteins may be recovered from CD membranes nearly quantitatively using denaturing detergent solutions such as 4 M guanidine hydrochloride/0.1% Triton X-100. Proteins can be fragmented using a variety of enzymes and proteases either directly on the Immobilon-CD membrane or after elution. If the proteins are to be digested after elution, the enzymes should be tested for compatibility with the elution buffer.

Proteins can also be eluted from Immobilon-P membranes, and two procedures for doing so are described in this unit: protein elution using detergents (Support Protocol 1; Szewczyk and Summers, 1988) and acidic elution with organic solvents (Support Protocol 2; D.W.S., unpub. observ.). Although gel electrophoresis combined with subsequent electroblotting is often the most powerful procedure for microscale protein isolation, identification, and characterization, one limitation of this system is that subsequent elution of proteins from PVDF membranes in good yields is often difficult.

Critical Parameters and Troubleshooting

The protocols listed in this unit should be sufficient for the transfer of most proteins. The efficiency of protein transfer is easily assessed by staining both the transfer membrane and the gel after transfer. When difficulty is encountered in the transfer recovery of a specific protein, several factors can be considered to improve transfer efficiency.

A first consideration is the type of gel used to separate proteins for transfer. Higher percentages of acrylamide make the transfer of proteins out of the gel more difficult. Gel thickness also affects transfer: the thicker the gel, the farther the protein has to migrate to reach the membrane. In general, effective transfer of proteins out of gels up to 1.5 mm thick can be achieved; thicker gels, although having increased protein capacity, offer little advantage owing to reduced removal of proteins from the thicker gel matrix.

In this unit, the transfer buffer described is essentially half-strength Towbin buffer (Towbin et al., 1979). This buffer contains a mixture of Tris and glycine and is the buffer of choice if the electroblotted proteins will be used for sequence analysis. Tris, glycine, and most other low-molecular-weight buffer components do not normally bind to PVDF membranes and can be completely removed by thorough rinsing with high-purity water. However, incomplete rinsing or partial drying of the membrane prior to rinsing can result in high residual levels of these compounds, which could be detrimental for sequence analysis.

The amino groups in the buffer act as scavengers for contaminants that could potentially modify proteins. Several simple precautions noted in Alternate Protocol 1 can be taken to minimize side reactions of proteins during electrophoresis and electroblotting (Speicher, 1989). If possible chemical modifications are not critical, several other buffers may be used. CAPS buffer (cyclohexylaminopropanesulfonic acid; LeGendre and Matsudaira, 1988) has a high pH (pH 11) and low ionic strength and, therefore, because of its low conductivity,

permits the use of high electric fields. For the transfer of very basic proteins that bind SDS poorly, high pH may prove advantageous in keeping the total net protein charge negative; however, the high pH of the CAPS buffer can also lead to deamination of sensitive side chains. Other types of electrotransfer buffers such as Tris/borate (Vandekerckhove et al., 1985) do not differ significantly in overall performance from the Tris/glycine buffer and may be used if N-terminal protein blockage is not a major concern.

Selection of membrane type is dependent on the ultimate use of the transferred protein, although high-retention PVDF membranes are most often chosen because they bind proteins more tightly than nitrocellulose membranes and have greater chemical and mechanical stability. For immunoblots, either nitrocellulose or Immobilon-P (Millipore) are preferred by some investigators because they produce a lighter background than high-retention PVDF membranes.

It is not uncommon in electroblot procedures to observe less than optimal transfer of proteins. If the membrane is entirely blank and little or no protein is left in the gel after transfer, it is likely that the electrodes were connected in a reversed order or that the membrane was placed on the wrong side of the gel. Blank spots on membranes are caused by air bubbles, which block the flow of current.

Another problem often encountered with poor transfer of proteins relates to the amount of SDS in the gel, which coats the proteins and improves mobility during electrotransfer. Nitrocellulose membranes are not sensitive to the amount of SDS in the gel and can even tolerate addition of SDS to the transfer buffer to encourage transfer of recalcitrant proteins. PVDF membranes are more sensitive to excess SDS, which can inhibit protein binding to the hydrophobic membrane; nevertheless, in a limited number of cases, addition of SDS to the transfer buffer may be beneficial. Low-retention PVDF membranes are particularly sensitive to excess SDS concentrations, whereas high-retention PVDF membranes are less so. Methanol is a common component of many transfer buffers because it facilitates dissociation of bound SDS from proteins. Therefore, if proteins transfer from the gel efficiently but do not bind well to PVDF membranes, the methanol concentration can be increased to 20% and/or the gel can be preequilibrated in transfer buffer for 15 to 30 min prior to transfer to reduce the SDS concentration (Mozdzanowski et al., 1992). Conversely, if proteins remain in the gel after transfer, the methanol concentration can be reduced or eliminated, and if necessary a low concentration of SDS (0.005%) can be added to the transfer buffer to facilitate electrotransfer from the gel.

If proteins have transferred efficiently but the stained membrane has blurred bands or a swirled pattern, it is likely that there was insufficient contact between the gel and membrane during transfer. The transfer sandwich should be held together firmly both to avoid having the gel shift position during transfer and to ensure close contact between the gel and membrane. Another common problem with transfers is the presence of bright white spots within protein bands on the membrane. These are due to air bubbles trapped between the gel and membrane or, in the case of PVDF membranes, possibly due to local drying of the membrane during transfer unit assembly. As indicated in the protocols, it is very important to check for air bubbles before electrotransfer. Air bubbles can be rolled out from the two surfaces by using a test tube, or the membrane can be carefully lifted to release the bubble and replaced without shifting the position of either gel or membrane.

Anticipated Results

Proteins in the middle two-thirds of the gel usually transfer to high-retention PVDF membranes with an average yield of 50% to 80%. The protein pattern on PVDF membranes stained with amido black or Coomassie blue should closely resemble the pattern on duplicate lanes of the gel stained with Coomassie blue. The staining intensity on the blot should be slightly higher than on the gel because the proteins are concentrated on the surface of the blot rather than distributed throughout the thickness of the gel. Proteins below the dye front usually will not be recovered on the membrane. Proteins in the top 20% of the gel are often incompletely transferred out of the gel.

Time Considerations

Assembling the transfer unit takes ~30 min. Electroblotting can be completed in 2 to 4 hr in most cases using a tank transfer unit and in ~1 hr using a semidry blotting unit.

BOSTON CENTRAL UNIVERSITY LIBRARY

Literature Cited

LeGendre, N. and Matsudaira, P. 1988. Direct protein microsequencing from Immobilon-P transfer membrane. *BioTechniques* 6:154-159.

Mozdzanowski, J., Hembach, P., and Speicher, D. 1992. High yield electroblotting onto polyvinylidene difluoride membranes from polyacrylamide gels. *Electrophoresis* 13:59-64.

Patterson, S.D., Hess, D., Yungwirth, T., and Aebersold, R. 1992. High-yield recovery of electroblotted proteins and cleavage fragments from a cationic polyvinylidene fluoride–based membrane. *Anal. Biochem.* 202:193-203.

Speicher, D.W. 1989. Microsequencing with PVDF membranes: Efficient electroblotting, direct protein adsorption and sequencer program modifications. *In* Techniques in Protein Chemistry (T. Hugli, ed.) pp. 24-35, Academic Press, San Diego.

Szewczyk, B. and Summers, D.F. 1988. Preparative elution of proteins blotted to Immobilon membranes. *Anal. Biochem.* 168:48-53.

Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Vandekerckhove, J., Bauw, G., Puypi, M., Van Damme, J., and Van Montagu, M. 1985. Protein-blotting on Polybrene-coated glass-fiber sheets. *Eur. J. Biochem.* 152:9-19.

Key References

LeGendre, N. and Matsudaira, P. 1988. See above.

Thoroughly reviews electroblotting using PVDF membranes.

Mozdzanowski, J. and Speicher, D.W. 1992. Microsequence analysis of electroblotted proteins. I. Comparison of electroblotting recoveries using different types of PVDF membranes. *Anal. Biochem.* 207:11-18.

Compares electroblotting recoveries using different PVDF membranes.

Contributed by Jeanine A. Ursitti,
Jacek Mozdzanowski, and
David W. Speicher
The Wistar Institute
Philadelphia, Pennsylvania

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Detection of Proteins on Blot Membranes

Staining of blot transfer membranes permits visualization of proteins and allows the extent of transfer to be monitored. In the protocols described in this unit, proteins are stained after electroblotting from one-dimensional or two-dimensional polyacrylamide gels to blot membranes such as polyvinylidene difluoride (PVDF), nitrocellulose, or nylon membranes (*UNIT 10.7*). PVDF is the preferred, more universal membrane and is emphasized here; however, most stains work similarly on nitrocellulose, and many can be used on alternative blotting membranes.

The first six Basic Protocols describe the use of six general protein stains—amido black, Coomassie blue, Ponceau S, colloidal gold, colloidal silver, and India ink. In addition, the fluorescent stains fluorescamine and IAEDANS, which covalently react with bound proteins, are described in Basic Protocol 7 and the Alternate Protocol. Table 10.8.1 lists approximate detection limits for each nonfluorescent stain as well as membrane compatibilities. The dimensional stability of blotted membranes facilitates direct, precise comparisons of staining patterns using different detection methods. For example, protein staining can be directly compared with immunoreactivity by staining one portion of a blot with a general protein stain while subjecting another portion containing duplicate samples to immunoblotting. When it is desirable to cut replicate lanes from the blot prior to any staining, prestained standards provide very useful visual reference points.

NOTE: High-purity water (from a Milli-Q purification system or equivalent) should be used throughout the protocols, and all plastic and glass boxes must be thoroughly cleaned before use to avoid staining artifacts. Membranes should be handled only by the edges with forceps. All steps should be performed at room temperature (unless otherwise described) and with gentle agitation. Use of an orbital shaker is recommended for steps that take longer than 1 min. If a PVDF membrane is allowed to dry after transfer, wet for 5 sec in 100% methanol and rinse with water before staining. Volumes of stain, destain, and wash solutions should be sufficient to cover the membrane and allow it to float freely. Unless noted otherwise, solutions may be stored for several months at room temperature.

Table 10.8.1 Staining Sensitivities and Membrane Compatibilities for Nonfluorescent Stains[a]

Stain	Minimum amount detected[b]	Membrane type[c]		
		PVDF	Nitrocellulose	Nylon
Amido black	50 ng	+	+	+
Coomassie blue	50 ng	+	–	+
Ponceau S	200 ng	+	+	+
Colloidal gold	2 ng	+	+	–
Colloidal silver	5 ng	+	+	–
India ink	5 ng	+	+	–

[a]Sensitivity of fluorescent stains is very dependent on protein amino acid composition.

[b]Minimum amount detected based on amount of protein loaded onto gel. The actual amount on the blot will be slightly lower because of losses during electrotransfer. Values are based on use of a full-sized gel (11 cm × 16 cm × 1.5 mm). Sensitivity will be ~2 to 5 times higher when minigels (8 cm × 10 cm × 1.0 mm) are used because the protein bands are concentrated on a smaller area of membrane.

[c]+ indicates stains well; – indicates membrane not compatible.

Copyright © 1995 by John Wiley & Sons, Inc.

BOWLING GREEN STATE UNIVERSITY LIBRARY

BASIC PROTOCOL 1

AMIDO BLACK STAINING

Amido black is used to stain proteins on blot transfer membranes. Transferred proteins (>50 ng/band) appear as dark blue bands on a light blue background. Amido black has a sensitivity similar to that of Coomassie blue, but it stains faster. It is the preferred stain for protein sequencing and in situ cleavage of proteins for determining internal sequences because the mild staining and destaining conditions minimize the likelihood that any protein will be extracted during the procedures.

Materials

Protein sample electroblotted to PVDF, nitrocellulose, or nylon membrane (*UNIT 10.7*)
Amido black 10B stain: 0.1% (w/v) amido black (naphthol blue black 10B, Sigma) in 10% (v/v) acetic acid
5% (v/v) acetic acid
Plastic boxes

1. Place blot transfer membrane in a plastic box and wash with water three times for 5 min each.
2. Stain membrane with amido black 10B stain for 1 min.

   Longer staining times only increase background and hence decrease sensitivity.
3. Destain the membrane with 5% acetic acid twice for 1 min each.
4. Rinse the membrane with water twice for 10 min each, then air dry.

BASIC PROTOCOL 2

COOMASSIE BLUE R-250 STAINING

Coomassie blue R-250 can be used with most types of blot membranes except nitrocellulose (high concentrations of organic solvents can dissolve nitrocellulose membranes). Coomassie blue has a similar sensitivity to amido black. Coomassie blue–stained proteins (>50 ng/band) appear as dark blue bands against a light blue background. The sequence of washing with water, staining, and destaining is similar to that for amido black staining (Basic Protocol 1), but the staining step is lengthened to 5 min.

Materials

Blot transfer membrane (*UNIT 10.7*)
Coomassie blue stain: 0.025% (w/v) Coomassie brilliant blue R-250 (Bio-Rad) in 40% methanol/7% acetic acid (v/v)
50% methanol/7% acetic acid (v/v)
Plastic box

1. Place blot transfer membrane in a clear plastic box. Wash with water three times for 5 min each.
2. Stain membrane with Coomassie blue stain for 5 min.

   If proteins on PVDF membranes are to be subjected to N-terminal sequencing, acetic acid should be omitted from the staining and destaining solutions to minimize protein extraction.
3. Destain membrane with 50% methanol/7% acetic acid for 5 to 10 min.
4. Rinse with water several times, then air dry.

PCU IASI / CENTRAL UNIVERSITY LIBRARY

PONCEAU S STAINING

BASIC PROTOCOL 3

Ponceau S is the least sensitive general protein stain described here. Transferred proteins (>200 ng/band) appear as red bands on a pink background. Major advantages of Ponceau S staining are that it is simple, rapid, and reversible. If desired, essentially all of the stain can be removed by extended destaining as described in steps 4 to 6. This can be particularly advantageous if the blot is to be reused after initial protein staining for a second detection method such as immunoblotting.

Materials

Blot transfer membrane (*UNIT 10.7*)
Ponceau S stain: 0.5% (w/v) Ponceau S (Sigma) in 1% (v/v) acetic acid
200 µM NaOH/20% (v/v) acetonitrile
Plastic box

1. Place blot transfer membrane in a plastic box and wash with water three times for 5 min each.
2. Stain membrane with Ponceau S stain for 30 sec to 1 min.
3. Destain membrane with several changes of water for 30 sec to 1 min each, then air dry. If the stain is to be extracted from protein bands (steps 4 to 6) prior to employing a second detection method, omit drying.

   Do not overdestain because the protein bands will be difficult to detect. Stop destaining when the background has a very slight pink tinge.

4. Make a permanent record of the staining pattern by photocopying or photographing the blot. Alternatively, mark the positions of the bands of interest directly on the blot by overlaying the wet blot with plastic wrap or a plastic bag and using a pencil or pen to outline the bands. Press with moderate pressure to make a permanent indentation on the membrane.
5. Extract Ponceau S stain from the protein bands with 200 µM NaOH/20% acetonitrile for 1 min.
6. Wash membrane with water three times for 5 min each, then air day.

COLLOIDAL GOLD STAINING

BASIC PROTOCOL 4

Colloidal gold is a highly sensitive stain. AuroDye (Amersham) is a ready-to-use commercial source of colloidal gold stain in a low-pH buffer. AuroDye should be used without dilution. Transferred proteins (>2 ng/band) will appear as red bands on a pink background. A higher signal may be obtained with alkali treatment of the membrane prior to staining by washing the membrane with 1% KOH followed by several rinses with phosphate-buffered saline (PBS). Glass rather than plastic boxes must be used to hold the membranes and solutions. Glass is easier to clean than plastic and is less likely to give artifacts. Ideally, a set of glass trays should be dedicated to the procedure.

Materials

Blot transfer membrane (*UNIT 10.7*)
Tween 20 solution: 0.3% (v/v) Tween 20 in PBS (*APPENDIX 2E*; prepare solution fresh weekly and store at 4°C)
AuroDye colloidal gold reagent (Amersham; store at 4°C)
Glass box

BOTANY / SOUTH TRAL UNIVERSITY LIBRARY

1. Place blot transfer membrane in a glass box. Wash with water three times for 5 min each.
2. Incubate membrane with Tween 20 solution for 30 min at 37°C with gentle agitation.
3. Wash membrane with Tween 20 solution three times for 5 min each at room temperature.
4. Rinse membrane several times with water.
5. Stain membrane with AuroDye for 2 to 6 hr (until desired color formation).

   Use only enough stain to cover the membrane completely.
6. Rinse membrane thoroughly with water, then air dry.

BASIC PROTOCOL 5

COLLOIDAL SILVER STAINING

Colloidal silver is a more economical stain than colloidal gold. The staining procedure is rapid, although the ferrous sulfate solution must be prepared immediately before use. Transferred proteins (>5 ng/band) appear as black bands on a light brown background for nitrocellulose and on a dark background for PVDF.

Materials

Blot transfer membrane (*UNIT 10.7*)
40% (w/v) sodium citrate (store at 4°C for several months)
20% (w/v) ferrous sulfate ($FeSO_4 \cdot 7H_2O$), prepared fresh
20% (w/v) silver nitrate (store at 4°C for several months)
Glass box

1. Place blot transfer membrane in a glass box and wash with water three times for 5 min each.
2. Add 5 ml of 40% sodium citrate and 4 ml of 20% ferrous sulfate to 90 ml water. Stir vigorously and slowly add 1 ml of 20% silver nitrate over ~1 to 2 min to form a suspension.
3. Immediately use the suspension to stain the membrane for ~5 min.

   The staining suspension should be used within 30 min.
4. Rinse membrane with water, then air dry.

BASIC PROTOCOL 6

INDIA INK STAINING

India ink is used to stain electroblotted proteins on blot transfer membranes. Transferred proteins (>5 ng/band) appear as black bands on a gray background. Sensitivity may be enhanced by brief alkali treatment of the membrane with 1% KOH followed by several rinses with PBS.

Materials

Blot transfer membrane (*UNIT 10.7*)
Tween 20 solution: 0.3% (v/v) Tween 20 in PBS (*APPENDIX 2E*; prepare solution fresh weekly and store at 4°C)
India ink solution: 0.1% (v/v) India ink (Pelikan 17 black) in Tween 20 solution (store 1 month at room temperature)
Plastic box

1. Place blot transfer membrane in a plastic box. Wash with water three times for 5 min each.

PCU IASI CENTRAL UNIVERSITY LIBRARY

2. Wash membrane with Tween 20 solution four times for 10 min each.
3. Stain membrane with India ink solution for 2 hr or overnight.
4. Rinse with water until an acceptable background is obtained, then air dry.

FLUORESCAMINE LABELING

BASIC PROTOCOL 7

Fluorescamine, or 4-phenylspiro[furan-2(3*H*),1′-phthalan]-3,3′-dione, is used to introduce a fluorescent label on electroblotted proteins via reaction with free amines. Transferred proteins are visualized on blot transfer membranes with UV light. This stain can be very sensitive and can be used in conjunction with a second detection method such as immunoblotting (also see Basic Protocol 3). However, the protein is irreversibly modified because fluorescamine reacts with available amino groups (i.e., lysines and the protein N terminus if it was not previously blocked).

Materials

Blot transfer membrane (*UNIT 10.7*)
Sodium bicarbonate solution: 100 mM sodium bicarbonate in 0.3% (v/v) Tween 20, pH 9.0 (prepare fresh weekly and store at 4°C)
Fluorescamine stain: 0.25 mg/ml fluorescamine (Sigma) in sodium bicarbonate solution (prepare fresh daily)
Plastic box

1. Place blot transfer membrane in a plastic box. Wash with water three times for 5 min each.
2. Wash membrane with sodium bicarbonate solution twice for 10 min each.
3. Label protein bands with fluorescamine stain for 15 min. Use enough staining solution to cover the membrane completely.
4. Wash membrane with bicarbonate solution three times for 5 min each.
5. Rinse membrane several times with water.
6. Visualize transferred proteins with UV light.

IAEDANS LABELING

ALTERNATE PROTOCOL

N-iodoacetyl-*N*′-(5-sulfo-1-naphthyl)ethylenediamine (IAEDANS or 1,5-I-AEDANS) is used for fluorescent labeling of electroblotted proteins on blot transfer membranes. Because IAEDANS reacts with free cysteines, disulfides in the sample must first be reduced with dithiothreitol (DTT). Transferred protein bands are visualized under UV light.

Additional Materials (also see Basic Protocol 7)

DTT solution: 200 mM dithiothreitol (DTT) in 100 mM Tris·Cl, pH 8.6 (*APPENDIX 2E*; prepare immediately before use)
100 mM Tris·Cl, pH 8.6 (*APPENDIX 2E*)
N-iodoacetyl-*N*′-(5-sulfo-1-naphthyl)ethylenediamine (IAEDANS; Sigma; store desiccated in the dark at –20°C)
Glass box

1. Place blot transfer membrane in a glass box. Wash with water three times for 5 min each.
2. Incubate membrane with DTT solution for 30 min to reduce disulfides in the sample.
3. Wash membrane with 100 mM Tris·Cl (pH 8.6) three times for 5 min each.

BOSTON CENTRAL UNIVERSITY LIBRARY

This washing step probably could be deleted if proteins have been electroblotted from a reducing gel (i.e., containing DTT and/or 2-mercaptoethanol). However, it is advisable to include this simple step routinely as some oxidation may occur during electrotransfer or on the blotted membrane itself during drying or storage.

4. Dissolve 86 mg IAEDANS (2 mM final concentration) in 100 ml of 100 mM Tris·Cl (pH 8.6).

   This step should be carried out in the dark (solution should be made fresh daily).

5. Add IAEDANS solution to the membrane and allow the reaction to proceed in the dark for 30 min with shaking.
6. Wash membrane with 100 mM Tris·Cl (pH 8.6) twice for 5 min each. Thoroughly rinse membrane with water to remove excess reagent.
7. Visualize the transferred proteins under UV light.

COMMENTARY

Background Information

The recent rapid expansion in the use of electrophoretic transfer of separated proteins to different types of membranes has necessitated the adaptation of existing protein staining techniques to transfer membranes. On-blot staining techniques serve multiple purposes, including detection of proteins for structural analysis and use in parallel with antibody reactivity to correlate precisely immunoreactivity with protein staining patterns. For the latter purpose an advantage of on-blot staining is that duplicate sections of membrane can be cut out and one section used for immunoblotting (*UNIT 10.10*) while a second section with duplicate lanes is stained with a general protein stain. The two pieces can then be precisely realigned. In contrast, a direct comparison of an immunoblot with a stained polyacrylamide gel is much less precise because of shrinking and swelling of the gel.

The most common membranes used for electroblotting are polyvinylidene difluoride (PVDF) and nitrocellulose. PVDF membranes have become increasingly popular because they are easy to handle and store, whereas nitrocellulose membranes are brittle and tend to break easily when dry. A number of PVDF membranes are now commercially available that have subtle but important differences in protein binding properties resulting from different proprietary manufacturing processes. For example, Bio-Rad Trans-Blot or Millipore Immobilon-P^{SQ} PVDF membranes generally show higher protein binding capacities and higher binding affinities compared to Immobilon-P (Mozdzanowski and Speicher, 1992). Use of high-retention PVDF membranes usually results in higher and more consistent electroblotting recoveries of most proteins than the use of either low-retention PVDF membranes or nitrocellulose. However, high-retention PVDF membranes tend to exhibit higher staining backgrounds, and it is more difficult to extract proteins or peptides from such membranes.

The protocols for staining with amido black, Coomassie blue, Ponceau S, and AuroDye follow the suppliers' recommendations. It should be noted that when staining PVDF membranes with Coomassie blue before N-terminal sequencing, omitting acetic acid from both the stain and destain solution is recommended to minimize potential extraction of protein from the membrane (Speicher, 1989).

The protocol for India ink staining of electroblotted proteins is essentially that of Hancock and Tsang (1983). Different brands of India ink may be used, but staining sensitivity may vary as a result.

The colloidal gold stain is the most sensitive membrane stain described here. As an alternative to commercially available colloidal gold stains, Moeremans et al. (1985) describe methods for preparing gold and iron solutions.

Fluorescent labels are advantageous because they can be used not only for sequential detection methods on the same blot with minimal potential interference, but also for detection prior to protein extraction from the membrane. For example, after visualization of proteins with a fluorescent label, the blot can be photographed and specific bands marked with a pencil, either directly on the membrane or through a plastic bag. The latter method leaves a permanent indentation on the membrane. The blot can then be probed with antisera (i.e., immunoblotted; *UNIT 10.10*).

PCU IASI-BEIJING UNIVERSITY LIBRARY

The protocol for fluorescamine labeling is based on the procedure described by Vera and Rivas (1988). Fluorescamine itself is not strongly fluorescent; however, when combined with the primary amines of proteins (i.e., N termini and lysine residues) it yields a highly fluorescent product. IAEDANS is an iodoacetic acid analog containing a naphthalene ring and is fluorescent under UV light. When the sample protein is reduced either prior to gel electrophoresis or with DTT after blotting, all cysteines are potentially available for reaction with IAEDANS. When the protein is not reduced, some cysteines may be involved in disulfide bonds and therefore not available for reaction. Additionally, some cysteines may be sterically inaccessible because of adsorption to the membrane and therefore will not react.

An additional visualization technique for PVDF membranes is transillumination, described by Reig and Klein (1988). In that technique, the membrane is dried at room temperature, then wet with 20% methanol and viewed on a white light box. Protein bands appear as clear areas. Sensitivity is usually comparable to that of Coomassie blue staining.

Critical Parameters and Troubleshooting

High-quality water (from a Milli-Q purification system or equivalent) should be used throughout these protocols. All plastic and glass boxes must be thoroughly cleaned by rinsing with water before use to avoid staining artifacts. Blot membranes should be handled by the edges only with gloves or, preferably, with forceps. This precaution is most critical for the more sensitive stains (i.e., colloidal gold, colloidal silver, and India ink). When using PVDF membranes, it is especially critical that the membrane does not dry between steps. If drying occurs, wet the PVDF membrane for 5 sec with 100% methanol, then rinse several times with water.

A brief alkali treatment can enhance staining with India ink or colloidal gold. In a procedure described by Sutherland and Skerritt (1986), the membrane is washed with 1% (w/v) KOH for 5 min followed by several rinses with PBS. The alkali treatment can easily be incorporated at the beginning of the procedures if desired.

Anticipated Results

Approximate detection limits and membrane compatibilities are listed in Table 10.8.1. It should be noted that detection limits may vary with gel size and the percentage of polyacrylamide in the gel used in preparation of the blot transfer membrane. The sensitivity of fluorescent stains is related to the number of reactive amino groups (fluorescamine) or cysteine residues (IAEDANS) present in the protein of interest.

Time Considerations

The total time required for staining with amido black, Coomassie blue, and Ponceau S is 30 min to 1 hr; colloidal gold requires 4 to 6 hr; colloidal silver requires 1 hr; whereas staining with India ink requires 2 hr to overnight. Fluorescent labeling requires ~1 hr. The optional alkali enhancement (see Critical Parameters and Troubleshooting) requires an additional 30 min at the beginning of the India ink and colloidal gold staining procedures.

Literature Cited

Hancock, K. and Tsang, V.C.M. 1983. India ink staining of proteins on nitrocellulose paper. *Anal. Biochem.* 133:157-162.

Moeremans, M., Daneels, G., and De Mey, J. 1985. Sensitive colloidal metal (gold or silver) staining of protein blots on nitrocellulose membranes. *Anal. Biochem.* 145:315-321.

Mozdzanowski, J. and Speicher, D.W. 1992. Microsequence analysis of electroblotted proteins. *Anal. Biochem.* 207:11-18.

Reig, J. and Klein, D. 1988. Submicrogram quantities of unstained proteins are visualized on polyvinylidene difluoride membranes by transillumination. *Appl. Theor. Electrophor.* 1:59-60.

Speicher, D.W. 1989. Microsequencing with PVDF membranes: Efficient electroblotting, direct protein adsorption and sequencer program modifications. *In* Techniques in Protein Chemistry (T.E. Hugli, ed.) pp. 24-35. Academic Press, San Diego.

Sutherland, M.W. and Skerritt, J.H. 1986. Alkali enhancement of protein staining on nitrocellulose. *Electrophoresis* 7:401-406.

Vera, J.C. and Rivas, C. 1988. Fluorescent labeling of nitrocellulose-bound proteins at the nanogram level without changes in immunoreactivity. *Anal. Biochem.* 173:399-404.

Key References

Moeremans, M., Daneels, G., and De Mey, J. 1985. See above.

Describes a method for preparation of colloidal metal stains.

Vera, J.C. and Rivas, C. 1988. See above.

Describes the use of multiple detection methods.

Contributed by Sandra Harper and David W. Speicher
The Wistar Institute
Philadephia, Pennsylvania

BAYLOR UNIVERSITY MEDICAL CENTER LIBRARY

Capillary Electrophoresis of Proteins and Peptides

Capillary electrophoresis (CE) is a high-resolution technique for the separation of a wide variety of molecules of biological interest such as metabolites, drugs, amino acids, nucleic acids, and carbohydrates. This unit focuses on the use of CE to separate proteins and peptides. As with polyacrylamide gel electrophoresis (PAGE; *UNIT 10.1*), CE separations of proteins and peptides are based on charge-to-mass ratios. While PAGE separations are restricted to polyacrylamide matrices and a relatively small number of buffer systems; CE separations can be achieved in a variety of different matrices using a wide range of electrophoresis buffers. Consequently, there is a much greater flexibility in the design of optimal separation protocols.

CE employs a fused-silica capillary column, which may or may not be derivatized, either in free solution or in the presence of a fluid matrix. In contrast to PAGE, the resulting protein or peptide bands cannot be fixed or stained. The CE separation is, in essence, a dynamic one and is more analogous to high-performance liquid chromatography (HPLC) than PAGE. However, CE separation is based on electrophoretic parameters, so it can be viewed as orthogonal to HPLC, which is based primarily on solvent partitioning principles. The combination of these two techniques is very powerful for analyzing and characterizing proteins and peptides. The initial expectation that CE would play a major role in protein separations has not been realized. However, it is particularly useful for determining the purity of a sample and for rapid, efficient, and quantitative evaluation of protein purification techniques.

For best separation results, it is necessary to optimize the procedure. If some properties of the protein are known—e.g., its isoelectric point (pI)—the selection of a suitable running buffer is relatively straightforward. In fact, CE itself can be used to determine the isoelectric point of a protein, either in purified form or in a mixture, by focusing the sample in a pH gradient that is generated within the capillary during electrophoresis (Basic Protocol 1). After refocusing, the sample is mobilized either by changing the anode buffer, or by applying a hydrodynamic force to the column to move the separated components past the detector. However, there is little information available that will predict the potential of a protein to interact with the capillary wall. Underivatized silica capillaries have a strong tendency to adsorb proteins, thereby immobilizing them and making protein separations impractical. If such interactions occur, then an additive or a coated capillary column is required. A certain amount of trial and error is involved, and it may be necessary to perform several CE runs to optimize the separation conditions for a given protein. One approach to optimization is described in Basic Protocol 2.

CE is most useful for separations of peptides (Basic Protocol 3) because it offers great flexibility in separation parameters. It can be used to monitor proteolytic digestions and optimize digestion conditions for the production of a representative peptide fingerprint of a protein. This profile can subsequently be used to provide structural information about the protein, especially when used in conjunction with reversed-phase HPLC (RP-HPLC). It can be used to screen peptide fractions that are obtained from a preparative RP-HPLC separation of a protease digestion. Fractions that contain single or major components are suitable candidates for protein sequence analysis. Similarly, CE can be used to assess the purity of synthetic peptides. In the presence of an internal standard, it can provide quantitative information about the various components that are present in a peptide mixture. CE can also be used as a micropreparative technique—with either multiple separations that are pooled (Basic Protocol 4) or a single, larger-scale separation (Alternate Protocol 1)—for the isolation of peptides from a protease digestion (in much the

Copyright © 1995 by John Wiley & Sons, Inc.

BOSTON CHURCH LIBRARY

same way that RP-HPLC is currently used). In most of these examples the same capillary column can be used for all the separations. Only changes in buffer composition, ionic strength, and the presence or absence of additives are required for each specific application.

INSTRUMENTATION

CE separation in its simplest form can be achieved by passing a high voltage between two buffer reservoirs that are joined by a liquid-filled fused silica capillary (Fig. 10.9.1). This results in the generation of electroosmotic flow (EOF; see Separation Theory), which allows the molecules of interest to be carried from one end of the capillary to the other. The capillaries are generally 30 to 50 cm long with 50 to 75 μm i.d. The net total volume of these capillaries is in the low microliter range. For comparison, the volume of a slab gel lane is ~1000 μl. The capillaries are thin walled; this allows rapid and efficient exchange of the Joule heating resulting from the high voltages (20 to 25 kV) that are necessary for electrophoretic separations. This heat dissipation minimizes the negative convective effects that could result in band broadening during electrophoresis. The fused-silica capillary is coated on the outside with a polyimide layer that confers excellent tensile strength to the otherwise fragile capillary. The polyimide sheathing is carefully burned from a small portion of the capillary to expose a section of the silica. This clear section of the capillary is inserted into the light path of a UV detector and becomes the flow cell. As the protein and peptide molecules are swept through the capillary by EOF, they pass through the detector light path and are registered on the UV monitor. In effect the capillary becomes a very-low-volume flow cell.

IMPORTANT NOTE: The removal of the polyimide coating makes the capillaries susceptible to breakage. Capillaries that are not provided in cartridges by the manufacturer should be handled with care to avoid breakage.

The combination of high field strength and large surface-to-volume ratios results in rapid and very efficient separations of both proteins and peptides. Sample loading volumes are routinely in the nanoliter range, with starting sample concentrations of ~0.1 μg/μl for UV

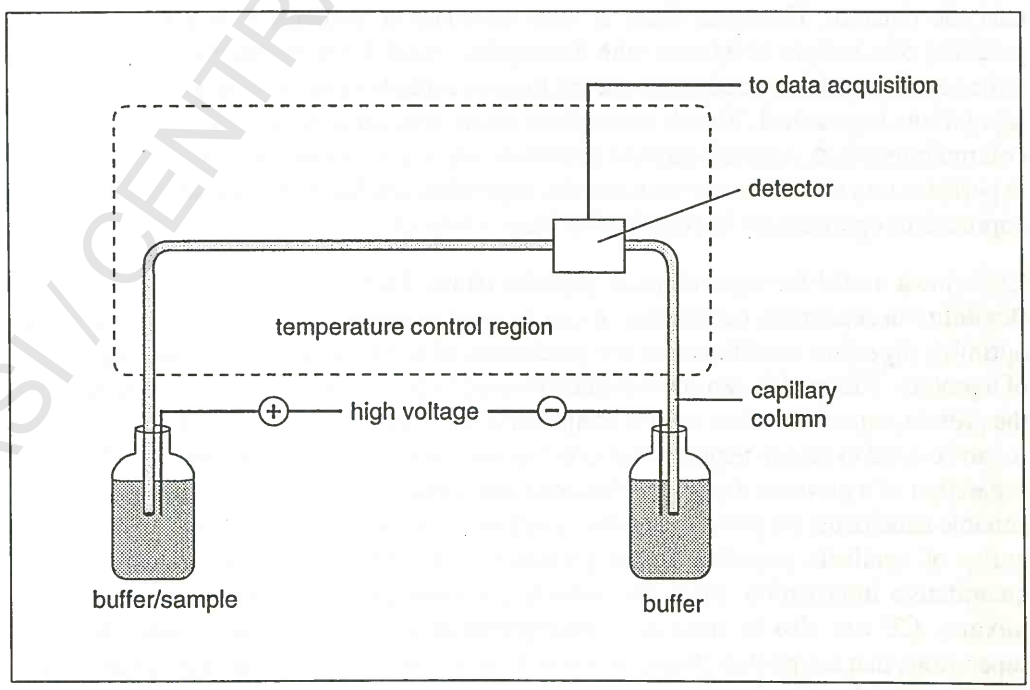


Figure 10.9.1 Schematic of a CE instrument.

EGU IASI / CENTRAL UNIVERSITY LIBRARY

Table 10.9.1 Commercially Available CE Instruments

Manufacturer	Model	Detection method	Reversal of polarity	Capillary heating	Cooling of sample/buffer	Capillary format	Software available
Applied Biosystems	270-HT	UV: 190-700 nm	Controlled by software	FAC	Sample	Free-hanging	DA
Beckman	P/ACE 5000 P/ACE 5510 P/ACE 5510	UV: filter; Diode-array: 190-600 nm; Fluorescence: argon laser; MS adapters	Manual	Liquid (Peltier)	Sample/buffer	Cartridge	C, DA
Bio-Rad	Biofocus 3000	UV: 190-800 nm	Controlled by software	Liquid	Sample/buffer	Cartridge	C, DA
Hewlett-Packard	HP 3D CE	Diode-array: 190-600 nm	Controlled by software	FAC (Peltier)	Sample/buffer	Cartridge	C, DA
Waters	Quantra 4000E	UV: filter	Manual	FAC (Peltier)	Sample/buffer	Free-hanging	DA

[a]All instruments have hydrostatic/electrokinetic sample loading and fraction collecting.

[b]Abbreviations: C, CE system control software; DA, data acquisition; FAC, forced-air convection.

detection. On-capillary preconcentration protocols are available when starting concentrations are below these levels. Clearly, with respect to sensitivity, speed, and versatility, CE can offer some significant advantages over gel electrophoresis for the separation of proteins and peptides.

As shown in Figure 10.9.1, CE can require minimal instrumentation. However, in reality, due to the high voltages that are utilized, safety issues require that the capillary column be incorporated into a dedicated CE instrument. These instruments can provide efficient capillary cooling and online detection of analytes. Table 10.9.1 lists a number of commercially available instruments and some of their characteristics. For the purpose of this discussion it is assumed that a generic CE system possessing the following capabilities is used: an active capillary cooling system, a UV detector capable of monitoring at 200 nm, a thermostatted autosampler, and a chromatographic data package. Given this basic unit a variety of separations can be performed, depending upon the nature of the proteins or peptides that are being separated.

SEPARATION THEORY

CE is part of the family of electrophoretic techniques that separate species based upon their size and ionic properties. An ion (i) placed into an electric field will move in the direction parallel to the field with a velocity (v_i) as follows:

$$v_i = \mu_i E$$

Here μ_i is the electrophoretic mobility of species i and E is the electric field, as defined by the following equation, where V is the voltage and L is the total length from electrode to electrode:

$$E = \frac{V}{L}$$

BCU ASTON CENTRAL UNIVERSITY LIBRARY

The electrophoretic mobility of an ion (μ_i) is:

$$\mu_i = \frac{q}{6 \pi \eta a_i}$$

where q is the charge on the ion, η is the viscosity of the solution, and a_i is the radius of the ion.

As seen from these three equations, the movement of the molecule in the column is dependent on the applied voltage, the length of the column, the charge on the molecule, and the size of the molecule.

In CE, buffer flow is generated inside the column when the electric field is applied. This flow is from the cathode electrode to the anode electrode in a fused-silica capillary column (from left to right in Fig. 10.9.1); this movement of the buffer is called the electroosmotic flow (EOF). In addition, normal electrophoretic separation occurs whereby a positively charged molecule moves in the same direction as the EOF, a negatively charged molecule moves in a direction opposite to the flow, and a neutral molecule is carried along by the flow. Thus, the total velocity of the molecule (μ_t) is given by the following equation, where μ_{eo} is the EOF in the column.

$$\mu_t = (\mu_{eo} + \mu_i) E$$

The magnitude of the EOF will vary as a function of the pH of the carrier buffer. In uncoated fused-silica columns, a low pH generates a slow EOF, whereas a high pH generates a fast flow. The flow can be suppressed or completely reversed depending on the coating on the column or organic modifier added to the carrier buffer (Landers et al., 1992; Tsuiji and Little, 1992). Thus, EOF is another parameter that can be optimized to aid in difficult separations. Neutral molecules, because they are all carried along by the flow, can be difficult to separate. Addition of a micelle to the carrier buffer results in the partitioning of molecules between the micelle and carrier buffer and can effectively resolve neutral molecules (Khaledi, 1994).

STRATEGIC PLANNING

Proteins can be separated on coated or uncoated capillary columns, and the choice of separation protocol depends on the specific properties of the target protein. The most important property is the pI, which can be determined by either conventional gel electrophoresis or CE. The use of a buffer ~2 pH units above the pI is optimal for CE separations. Table 10.9.2 lists a number of buffers and their pH ranges. The presence of high salt concentrations in the sample can interfere with the separation process, so dialysis or dilution may be required. Protein concentrations of microgram per microliter are optimal for CE separations; however, on-capillary concentration protocols exist that allow the separation of proteins with concentrations 1 to 2 orders of magnitude lower.

Peptides can be effectively separated in open-tube fused-silica capillary columns. The EOF generated within the capillary causes separation of both charged and neutral peptides. The respective migration times are dependent upon both the pH of the electrophoretic separation and the charge-to-mass ratios of the peptides. For example, at low pH, peptides with net positive charge migrate towards the anode faster than neutral or negatively charged peptides. Even negatively charged peptides are swept towards the anode by EOF. In practice this migration order is further modified by the mass of the peptides, with small positively charged peptides having faster migration rates.

The choice of buffer for CE is of primary importance. The use of buffer systems that separate at high pH significantly alters the migration positions of peptides relative to those

PGU IAS / CENTRAL UNIVERSITY LIBRARY

Table 10.9.2 Useful Electrolytes for CE of Peptides

Electrolyte	Useful pH range	Minimum useful wavelength (nm)
Phosphate	1.14-3.14	195
Citrate	3.06-5.40	260
Acetate	3.76-5.76	220
MES	5.15-7.15	230
Phosphate	6.20-8.20	195
Tris	7.30-9.30	220
Borate	8.14-10.14	180

of low-pH separations. The selection of a specific buffer is dictated both by its buffering capacity at a selected pH and by its minimum usable UV absorbance wavelength. Characteristics of some useful buffer systems are shown in Table 10.9.2. Small differences in the pH or composition of the buffer can have a significant impact on the absolute mobilities of the peptides, so it is advisable to adjust the pH of buffers accurately and reproducibly. Alternatively, a suitable internal standard can be incorporated into the separation in order to obtain relative mobilities. For the separation of peptides from a tryptic digest of a protein, where almost every peptide carries at least two positive charges, a phosphate buffer of pH 2.0 to 3.0 is commonly used. Likewise, for V8 protease digestions (which cleaves C-terminal to glutamic acid residues), the preferred separation conditions use a borate buffer with a pH range of 8.0 to 9.0.

SEPARATION OF PROTEINS BY ISOELECTRIC FOCUSING

BASIC PROTOCOL 1

Isoelectric focusing (IEF) separations can be performed readily by CE; this can be a useful first step in selecting subsequent buffer systems for protein separations. Proteins and peptides are amphoteric in nature, so their charge is dictated by the surrounding carrier buffer. When a protein or peptide is placed in an electric field, it moves to a region where the surrounding pH equals its isoelectric point (pI); at this point the molecule stops moving. A pH gradient is generated in a coated capillary column by filling the column with a sample solution that contains ampholytes. A high-pH solution (sodium hydroxide) is placed in the cathode reservoir and a low-pH solution (phosphoric acid) is placed in the anode reservoir. An electric field is then applied across the column and the ampholytes and the sample move to a location in the column corresponding to their respective pIs. If the molecule drifts out of the isoelectric region, a charge is induced by the surrounding pH solution and the molecule moves back to the region of zero charge. Narrow regions are formed with pI resolutions on the order of 0.01 pH units. To determine the pI location of a particular protein in the column, markers of known pIs are added to the sample mixture, and extrapolation between markers yields the pI of the protein of interest.

After separation of the sample in the absence of EOF, mobilization of the proteins past the detector is necessary. A simple replacement of the anode reservoir with the buffer in the cathode reservoir establishes EOF and causes the sample zone to migrate pass the detector. The separation discussed here assumes there is no EOF in the column during focusing. However, separations may also be done in the presence of EOF (Mazzeo and Krull, 1992).

BOURNEMOUTH UNIVERSITY LIBRARY

Materials

Ampholyte mixture, pH 3 to 10 (Bio-Rad)
Sample containing 0.5 to 1.0 mg/ml protein
IEF markers (optional)
20 mM sodium hydroxide (NaOH; store at 4°C)
10 mM phosphoric acid (store at 4°C)

50-µm-i.d. coated silica capillary column
CE instrument (see Table 10.9.1)

1. Add ampholyte mixture to 0.5 ml protein sample to give a final concentration of 2.5% (v/v) ampholytes.

   If desired, add IEF markers to a final concentration of 0.1 mg/ml to calibrate the column.

2. Add the solution to the sample reservoir. Fill the capillary by pressurizing the reservoir (0.5 psi).

3. Place 10 mM phosphoric acid in the anode reservoir (negative end). Place 20 mM NaOH in the cathode reservoir (positive end).

4. Focus the sample 4 to 6 min at 8 to 10 kV constant voltage. Monitor the current until it reaches steady state.

   The instrument should be operated in accordance with the manufacturer's instructions.

5. Mobilize the sample by placing 20 mM NaOH in the anode reservoir and setting the voltage to 10 kV. Monitor the mobilization of the proteins past the detector 15 to 20 min.

   Another method for mobilizing the proteins past the detector is to apply a hydrodynamic force to one end of the column and push the zones past the detector. Commercially available CE instruments have this capacity as an option. To minimize distortion of the pI regions, a constant electric field (10 kV) must be applied at the same time. The pressure applied is small (0.5 psi) and is maintained for the entire time of mobilization.

6. Wash the column after each run with 10 mM phosphoric acid for 1 min at 0.5 psi. Store the column at room temperature in running buffer (short-term) or water (long-term).

BASIC PROTOCOL 2

SEPARATION OF PROTEINS

To use CE for protein analyses, the more that is known about the particular protein, the higher the success rate for developing a separation method. Knowing the pI of a protein permits selection of the appropriate run buffer to maximize differences in the charge-to-size ratio. IEF separation performed using CE (Basic Protocol 1) can be a useful first step in selecting subsequent buffer systems for protein separations on an underivatized capillary column. If no protein is detected, it has probably adsorbed to the silica surface of the column. Repeating the separation in the presence of an ionic detergent such as SDS will coat the protein with negative charge and prevent adsorption. However, this means that proteins will separate on the basis of size alone. Alternatively, if the pI of the protein is unknown, the separation can often be achieved by using a coated capillary column in the presence of high-pH, high-ionic-strength buffer.

RCU IASVIC CENTRAL UNIVERSITY LIBRARY

Materials

Sample containing 10 mg/ml protein
50 mM and 500 mM sodium borate, pH 8.0 to 9.5 (for separation of proteins with unknown pI)
5 mM and 50 mM buffer at pH >pI (for separation of proteins with known pI; see Table 10.9.2)
50-µm-i.d. coated (unknown pI) *or* uncoated (known pI) fused-silica capillary column
CE instrument (see Table 10.9.1)

For separations of proteins with unknown pI

1a. Dilute protein sample 1:10 (v/v) with 50 mM sodium borate buffer to give a final concentration of 1.0 mg/ml.

Alternatively, dialyze sample (at a concentration of 1 mg/ml) in 50 mM sodium borate buffer.

2a. Fill a coated column with 500 mM sodium borate buffer.

3a. Inject the sample hydrodynamically, e.g., 3 sec at 1 psi.

Consult the manufacturer's instruction manual for the proper procedure.

4a. Separate using the following conditions:

Detector wavelength: 200 nm
Temperature: 25°C
Run voltage: 10 kV
Run time: 30 min.

For separations of proteins with known pI

1b. Dilute protein sample 1:10 (v/v) with a 5 mM buffer that has a pH above the pI of the protein to induce a charge on the protein (final protein concentration = 1.0 mg/ml).

Alternatively, dialyze sample (at a concentration of 1 mg/ml) in 5 mM buffer.

2b. Fill an uncoated capillary column with the same buffer at 50 mM.

3b. Inject the sample hydrodynamically, e.g., 3 sec at 1 psi.

Consult the manufacturer's instruction manual for the proper procedure.

4b. Separate using the following conditions:

Detector wavelength: 200 nm
Temperature: 25°C
Run voltage: 25 kV
Run time: 30 min.

If the protein of interest is not detected, add 10 mM SDS to the run buffer to induce a charge on the molecule. If the peaks are too broad, add 0.1% (w/v) methylcellulose to coat the column and reduce adsorption of the molecule onto the wall of the column. The addition of SDS and/or methylcellulose will increase the run time of the separation.

BCU IALH CENTRAL UNIVERSITY LIBRARY

ANALYTICAL PEPTIDE SEPARATIONS

CE is particularly useful for separating specific peptides from complex mixtures such as the products of proteolytic digestion. In this procedure, the peptide mixture to be analyzed is placed in a sample vial at the cathode end of the capillary column. The anode reservoir contains running buffer. The sample is loaded into the capillary either electrokinetically or by pressure or vacuum, then the sample vial is replaced by running buffer and electrophoresis begins. The following protocol describes the use of a P/ACE 5000 CE instrument (Beckman) using low-pressure sample injection. However, other instruments (see Table 10.9.1) are capable of similar separations when operated according to the manufacturer's instructions.

Materials

Peptide mixture: e.g., tryptic digest of β-lactoglobulin (Applied Biosystems)
0.05 M and 0.25 M sodium phosphate buffer, pH 2.30 (store at 4°C)
0.1 M sodium hydroxide (NaOH)

75-μm-i.d. fused-silica capillary column (Beckman)
CE instrument (e.g., P/ACE 5000, Beckman, or equivalent; see Table 10.9.1)

1. Precondition the capillary by flushing with the following solutions:

   10 column volumes of 100 mM NaOH at low pressure (0.5 psi)
   10 column volumes water
   4 column volumes of 0.25 M sodium phosphate buffer, pH 2.3.

   Store the column in 0.25 M sodium phosphate buffer, pH 2.3, at 25°C.

   When changing to a different separation buffer, equilibrate the column 4 hr with the new buffer (Strickland and Strickland, 1990).

   If a new capillary is being used, it is essential that this preconditioning step be performed prior to attempting separations.

2. Prepare the peptide mixture by dissolving 10 nmol (~300 μg) in 10 ml of 0.05 M sodium phosphate buffer, pH 2.30. Freeze unused mixture in 100-μl aliquots.

3. Apply 10 to 20 nl sample using low-pressure injection for 10 sec at 0.5 psi.

   Consult the manufacturer's manual for the proper procedure.

4. Separate the peptide mixture using the following conditions:

   Electrolyte: 0.05 M sodium phosphate buffer, pH 2.3
   Detector wavelength: 200 nm
   Temperature: 25°C
   Voltage: 25 kV.

   Consult the manufacturer's manual for proper operating conditions.

5. Rinse the column with the following solutions:

   0.5 min with water
   1.0 min with 0.1 N NaOH
   1.5 min with water
   1 min with 0.25 M sodium phosphate buffer, pH 2.30.

   Store the column in running buffer or water at room temperature.

   An example of a separation protocol for a Beckman P/ACE 5000 instrument with a 37-min run time is described in Table 10.9.3.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Table 10.9.3 Operating Program for Peptide Analysis[a]

Step	Process	Duration	Inlet[b]	Outlet[b]	Control summary[c]
1	Set Temp	—	—	—	Temp 25°C
2	Set Detector	—	—	—	UV 200nm
3	Rinse	1.0 min	33	7	Forward: HP
4	Rinse	1.5 min	29	7	Forward: HP
5	Rinse	0.5 min	29	9	Forward: HP
6	Inject	10.0 sec	11	9	Low pressure
7	Separate	30.0 min	29	9	Voltage 25.00 kV
8	Rinse	0.5 min	33	7	Forward: HP
9	Rinse	1.0 min	32	7	Forward: HP
10	Rinse	1.5 min	33	7	Forward: HP
11	Rinse	1.0 min	34	9	Forward: HP

[a]For analysis of proteolytic digests, HPLC fractions, and synthetic peptides.

[b]Vial contents: 7, waste vial (water level just to contact capillary effluent); 9, electrolyte (50 mM sodium phosphate buffer, pH 2.30); 11, sample; 29, electrolyte (50 mM sodium phosphate buffer, pH 2.30); 32, regeneration solution (0.1 N NaOH); 33, water; and 34, 0.25 H phosphate buffer, pH 2.30.

[c]Abbreviation: HP, high pressure.

B. C. U. „M. EMINESCU" IAŞI
CHIMIE

MICROPREPARATIVE CAPILLARY ELECTROPHORESIS: MULTIPLE SEPARATIONS

BASIC PROTOCOL 4

Although CE has been most commonly used for analytical separations, considerable interest has developed in using its high resolving power in micropreparative applications. Two basic approaches have evolved, one utilizing multiple separations and collections from an analytical capillary (Bergman and Jornvall, 1992) and the other utilizing a single separation and collection on a much larger-diameter (150-µm-i.d.) capillary (see Alternate Protocol). For the multiple collection approach to be effective, the elution times of the peptides must be reproducible. The following method uses the P/ACE 5000 capillary electrophoresis instrument to separate a mixture of peptides.

Materials

Peptides: ACTH 4-100, angiotensin I, and angiotensin II (Sigma)
0.05 mM and 0.25 mM sodium phosphate buffer, pH 2.30 (store at 4°C)
0.1 M sodium hydroxide (NaOH)

75-µm-i.d. fused-silica capillary column (Beckman)
CE instrument (e.g., P/ACE 5000, Beckman, or equivalent; see Table 10.9.1)
Conical microvials (Beckman)

1. Precondition the capillary column by flushing with the following solutions at low pressure (0.5 psi):

   10 column volumes of 0.1 M NaOH
   10 column volumes of water
   4 column volumes of 0.25 M sodium phosphate buffer, pH 2.3.

   Store in 0.25 M sodium phosphate, pH 2.30.

   When changing to a different separation buffer, equilibrate the column 4 hr with the new buffer (Strickland and Strickland, 1990).

2. Prepare a peptide mixture by dissolving 0.2 mg each of ACTH 4-10, angiotensin I, and angiotensin II in 1 ml of 0.05 M sodium phosphate buffer, pH 2.3 (final concentration = 0.6 mg/ml peptide). Store 100-µl aliquots at –20°C.

BCU IAŞI / COLECŢIA CENTRALĂ UNIVERSITY LIBRARY

3. Load 10 to 20 nl peptide mixture by low-pressure injection for 10 sec at 0.5 psi.

Consult the manufacturer's manual for the proper procedure.

4. Separate the mixture using the following conditions:

Electrolyte: 0.05 M sodium phosphate, pH 2.30
Detector wavelength: 200 nm
Temperature: 25°C
Run voltage: 25 kV
Fraction size: 1 min.

Consult the manufacturer's manual for proper operating conditions.

5. Replace the standard outlet reservoir (anode) with a series of conical microvials, each containing 10 µl of 0.05 M sodium phosphate buffer, pH 2.30. Collect 3-min fractions into each of the vials for the length of the separation.

6. Repeat the injection and separation (steps 3 to 5) four times and combine the contents from the corresponding fraction numbers.

It will take 2 to 3 hr to complete the fractionation.

7. Screen fractions for peptide content using the analytical separation previously described (see Basic Protocol 3).

ALTERNATE PROTOCOL

MICROPREPARATIVE CAPILLARY ELECTROPHORESIS: SINGLE SEPARATION

In this micropreparative protocol, a larger quantity of the peptide mixture is loaded onto a 150-µm-i.d. capillary column and the fractions are collected from a single separation (Kenny et al., 1993). For the single-collection approach the CE instrument must be able to effectively control the capillary temperature at the elevated power levels that are required for electrophoresis. Forced air cooling is generally inadequate for this purpose.

***Additional Materials** (also see Basic Protocol 4)*

4:1 (v/v) 0.5 M sodium phosphate buffer (pH 2.50)/ethylene glycol
150-µm-i.d. fused-silica capillary column (Polymicro)

1. Prepare peptide mixture and precondition the column (see Basic Protocol 4, steps 1 and 2).

2. Load 0.1 µl peptide mixture by low-pressure injection for 10 sec at 0.5 psi.

Consult the manufacturer's manual for the proper procedure.

3. Separate the peptide mixture using the following conditions:

Electrolyte: 0.05 M sodium phosphate buffer, pH 2.30
Detector wavelength: 200 nm
Temperature: 25°C
Run voltage: 7.5 kV
Fraction size: 3-min per collection vial.

Consult the manufacturer's manual for proper operating conditions.

4. Replace the standard outlet reservoir (anode) by a series of conical microvials each containing 10 µl of 4:1 (v/v) 0.5 M sodium phosphate buffer/ethylene glycol. Collect 3-min fractions into each vial for the length of the separation.

It will take ~2 hr to fractionate the sample.

FCU IASI / CENTRAL UNIVERSITY LIBRARY

5. Screen fractions for peptide content using the analytical separation protocol (see Basic Protocol 3).

COMMENTARY

Background Information

Protein separation is one of the more difficult types of separation to perform by capillary electrophoresis (CE; Novotny et al., 1990). Each protein has it own set of optimal separation conditions, and one optimized protocol will not necessarily transfer well to the separation of a different protein. The more that is known about the properties of the molecule, the easier it will be to optimize the separation conditions. The most important property is probably the pI, because it will dictate the selection of pH and ionic strength conditions for the initial separation attempt. However, a variety of other properties must also be taken into account when optimizing a separation protocol: shape (globular or fibrous), aggregation tendencies, solubility in dilute solutions, salt requirements, thermal stability, and hydrophobic nature. Several parameters can be varied within CE separations in order to utilize this knowledge—e.g., temperature, buffer/salt selection, ionic and nonionic detergents, and matrix additives (polyacrylamide, methylcellulose). Several protocols have been developed that can serve as useful starting points for most protein separations. Table 10.9.4 lists references for some groups of proteins. Furthermore, many of the manufacturers listed in Table 10.9.1 have developed additional protocols for protein separations.

Ultimately, the selection of an appropriate separation protocol for a protein depends on the specific properties of that protein. However, there are separation approaches that utilize specific properties of proteins. IEF-CE separates on the basis of charge alone; CE in the presence of ionic detergents such as SDS separates on the basis of size alone; and CE using underivatized capillaries at low pH or derivatized capillaries at high salt and high pH separates on the basis of both size and charge. Successful CE separations also depend on the behavior of other contaminants that are present in the mixture. It may also be possible to take advantage of the knowledge gained from previous purification steps, e.g., ion-exchange or size-exclusion chromatography, to select CE separation conditions.

Analytical-scale separation of peptides by CE has found application in a variety of areas. Proteolytic digestions can be monitored by placing the digestion mixture in a sample vial on the sample table that is equilibrated to 37°C. The instrument can then be programmed to analyze an aliquot at desired intervals, such as every 4 hr. Digestion is complete when a stable CE profile is obtained. Also, fractions collected from an HPLC separation of a protease digest can readily be screened orthogonally by CE to determine the homogeneity of any given peak. Fractions with a single major component can then be subjected to protein sequence analysis.

The sensitivity of UV detection at 200 nm is such that digest levels of ≥100 pmol can be effectively screened by this method. Below these load levels the fractions have to be reduced in volume or preconcentrated on the capillary to achieve a sufficiently high concentration for visualization (Dolnik et al., 1990; Burgi and Chien, 1992).

Selection of an appropriate micropreparative protocol will depend on the type of instrumentation that is available. The single-separation, large-diameter capacity protocol will only work with a CE instrument that has adequate cooling capabilities. The multiple-separation, analytical capillary protocol will work on any CE instrument that has reproducible and stable electrophoretic mobilities.

Critical Parameters

For protein analysis, the charge on the protein is the most important parameter for sepa-

Table 10.9.4 References for Separations of Different Protein Types

Protein type	Reference
Basic proteins	Wiktorowicz (1990); Bullock (1991)
Glycoproteins	Tran et al. (1991); Tsuiji (1992)
Antibody-antigen complexes	Nielsen et al. (1991)
Milk proteins	Chen et al. (1992)

BCU SHORT TERM LOAN LIBRARY

ration. However, the charge-to-shape ratio also plays a role in effective separation of proteins. The charge on the protein is controlled by the buffer and the additives used for the run. The shape is determined by the denaturing conditions of the solutions. The surface charge on a protein may ultimately dictate the extent of wall interactions in addition to electrophoretic mobility, whereas shape has its greatest effect upon mobility alone.

The recovery of peptides from underivatized silica capillary columns appears to be very good—even at subpicomole levels. However, proteins are frequently irreversibly bound to these same capillaries. Clearly, at some point a "peptide" must become a "protein" in the context of CE. There is little published data available to clarify this issue, although good recoveries of peptides 30-40 amino acids in length have been obtained from underivatized capillary columns. If the recoveries of higher molecular weight peptides are poor, then the separation should be repeated using a coated capillary column. During peptide synthesis, single-amino-acid-deletion products are frequently produced. The properties of these deletion products are very similar to those of the expected species, so the deletion products may not be readily resolved by reversed-phase HPLC. In this case, CE has proven to be a very useful orthogonal technique because of its high resolution power and separation based on charge-to-mass ratios. However, it must be emphasized that the effective separation and detection of proteins and peptides requires relatively high solute starting concentrations.

When using uncoated columns, a sodium hydroxide rinse is required to maintain the capillary surface. However, this rinse is never used with coated capillaries because it will remove the coating. Commercially available coated columns are satisfactory but linear polymers (e.g., methylcellulose) have been used with more success (Palmieri and Nolan, 1994). If a protocol can be worked out for the molecule of interest, rapid CE analysis with little loss of starting material can augment the tools used now for protein analysis (e.g., HPLC).

Troubleshooting

Mechanical and electrical problems that might be encountered during instrument use are addressed in the troubleshooting sections of the manufacturers' manuals. One of the common difficulties is the loss of electrical contact during a run. This loss of current can be caused if a small bubble forms in the column during injection or if the power generated by the run is so great that the solution boils or outgasses. Purging the column after a failed run removes any bubbles in the column. The sample vial must contain enough sample to cover the end of the column during injection to prevent injection of air into the column. The sample vial can dry out during a long run, so sample temperature control or the use of the proper cap on the vial is required to slow evaporation. Reducing the buffer concentration or the run voltage can eliminate bubble formation during a run. Degassing the run buffer is also useful if the outgassing problems continue.

The instrument sometimes arcs during a run; this is caused by salt deposits or moisture on the high-voltage electrode stations. Particular attention should be given to keeping the electrode stations salt free. Moisture buildup can be minimized by reducing the separation voltage by 5 kV. Frequent visual inspection is also advised.

The column can plug up because of salt deposition if the column dries out. Washing the column with water dissolves the plug; storing the column in water if it is not going to be used for a long period of time prevents plugging. Proteins can precipitate during a run and plug the column. Using lower protein concentrations prevents this type of plugging.

Anticipated Results

Electrophoretic profiles for both protein and peptide separations should be highly reproducible. However, absolute migration times may vary from run to run, especially if the separation temperature is not adequately controlled. For this reason, the use of an appropriate internal standard is highly recommended (see the manufacturer's instructions).

At load levels <50 pmol of digest, the recovery of peptides from a narrow-bore (1-mm) HPLC separation is poor. A convenient upper separation range for micropreparative CE is ~50 pmol of digest. The use of CE as a preparative separation technique therefore complements conventional HPLC capabilities. Recoveries from micropreparative CE separations are essentially quantitative down to the low picomole level, although postseparation handling losses can be significant. To use the multiple separation approach, the CE instrument must be capable of maintaining very stable migration times. Not all instruments have this capability, so it is advisable that the method be used for simple mixtures of peptides. The single-separation approach is recommended for more com-

plex mixtures. Because the fractions are collected "blind," they should be screened by matrix-assisted laser desorption/ionization (MALDI; *UNIT 11.6*) mass spectrometry prior to sequencing.

Time Considerations

Complete analysis times for IEF are ~30 min. Analysis times for protein separations are ~30 min for each step in method development. Peptide separations are normally completed in 40 min and frequently in as little as 20 min. Micropreparative peptide separations using either the multiple or single-separation approach are normally achievable within 90 min, but may take 2 hr or longer. It should be noted that these times are approximate, and sufficient time should be allowed to electrophorese all the components completely from the capillary.

Literature Cited

Bergman, T. and Jornvall, H. 1992. Capillary electrophoresis for preparation of peptides and direct determination of amino acids. *In* Techniques in Protein Chemistry III (R.A. Hogue Angeletti, ed.) pp. 129-134. Academic Press, San Diego.

Bullock, J.A. and Yuan, L.-C. 1991. FSCE of basic proteins in uncoated fused-silica capillary tubing. *J. Microcol. Sep.* 3:241-248.

Burgi, D.S. and Chien, R.-L. 1992. On-column sample concentration using field-amplification in CZE. *Anal. Chem.* 64:489A-495A.

Chen, F.A., Kelly, L., Palmieri, R., Biehler, R., and Schwartz, H. 1992. Use of high ionic strength buffers for the separation of proteins and peptides with CE. *J. Liq. Chromatogr.* 15:1143-1161.

Dolnik, V., Cobb, K.A., and Novontny, M. 1990. Capillary zone electrophoresis of dilute samples with isotachophoretic preconcentration. *J. Microcol. Sep.* 2:127-140.

Kenny, J.W., Ohms, J.I., and Smith, A.J. 1993. Micropreparative capillary electrophoresis (MPCE) and micropreparative HPLC of protein digests. *In* Techniques in Protein Chemistry, IV (R.A. Hogue Angeletti, ed.) pp. 363-370. Academic Press, San Diego.

Khaledi, M.G. 1994. Micellar electrokinetic capillary chromatography. *In* Handbook of Capillary Electrophoresis (J.P. Landers, ed.) pp. 43-94. CRC Press, Boca Raton, Fla.

Landers, J.P., Madden, B.J., Oda, R.P., and Spelsberg, T.C. 1992. The use of modifiers of endoosmotic flow for analysis of microheterogeneity. *Anal. Biochem.* 205:115-124.

Mazzeo, J.R. and Krull, I.S. 1992. Improvements in the methods developed for performing isoelectric focusing in uncoated capillaries. *J. Chromatogr.* 606:291-296.

Nielsen, R.G., Richard, E.C., Santa, P.F., Sharknas, D.A., and Sittampalam, G.S. 1991. Separation of antibody-antigen complexes by CZE, IEF, and HPSEC. *J. Chromatogr.* 539:177-185.

Novotny, M.V., Cobb, K.A., and Liu, J. 1990. Recent advances in CE of proteins, peptides, and amino acids. *Electrophoresis* 11:735-749.

Palmieri, R. and Nolan, J.A. 1994. Protein capillary electrophoresis: Theoretical and experimental considerations for methods development. *In* Handbook of Capillary Electrophoresis (J.P. Landers, ed.) pp. 325-368. CRC Press, Boca Raton, Fla.

Strickland, M. and Strickland, N. 1990. FSCE using phosphate buffer and acidic pH. *Am. Lab.* 22:60-65.

Tran, A.D., Parks, S., Lisi, P.J., Huynh, O.T., Ryall, R.R., and Lane, P.A. 1991. Separations of carbohydrate-mediated microheterogeneity of human erythropoietin by free solution CE. *J. Chromatogr.* 542:459-471.

Tsuiji, K. and Little, R.J. 1992. Charge-reversed, polymer-coated capillary column for the analysis of a recombinant chimeric glycoprotein. *J. Chromatogr.* 594:317-324.

Wiktorowicz, J.E. and Colburn, J.C. 1990. Separation of cationic proteins via charge reversal in CE. *Electrophoresis* 11:769-773.

Key References

McCormick, R.M. 1994. Capillary zone electrophoresis of peptides. *In* Handbook of Capillary Electrophoresis (J.P. Landers, ed.) pp. 287-324. CRC Press, Boca Raton, Fla.

Excellent review of both theory and practice for separation of peptides by CE.

Palmeiri and Nolan, 1994. See above.

Good review of methods development for protein separations.

Contributed by Dean Burgi
Genomyx
Foster City, California

Alan J. Smith
Stanford University Medical Center
Stanford, California

BCU CENTRAL UNIVERSITY LIBRARY

Immunoblot Detection

Immunoblotting (often referred to as western blotting) is used to identify specific antigens recognized by polyclonal or monoclonal antibodies. Protein samples are solubilized, usually with sodium dodecyl sulfate (SDS) and in selected cases with reducing agents such as dithiothreitol (DTT) or 2-mercaptoethanol (2-ME); some antibody epitopes are destroyed if reducing conditions are used. Following solubilization, the material is separated by SDS-PAGE (using either one- or two-dimensional gels; *UNITS 10.1 & 10.4*). The antigens are then electrophoretically transferred in a tank or a semidry electroblotting unit to a nitrocellulose, polyvinylidene difluoride (PVDF), or nylon membrane (*UNIT 10.7*). When nitrocellulose or PVDF membranes are used (*UNIT 10.8*), the process can be monitored by a reversible staining procedure with Ponceau S. After staining, protein bands on the membrane can be photographed and/or the positions of the detected proteins can be marked with indelible ink (e.g., Paper-Mate pen). The membrane is then completely destained by soaking in water for an additional 10 min.

At this point the transferred proteins are bound to the surface of the membrane, providing access for reaction with immunodetection reagents. All remaining binding sites are blocked by immersing the membrane in a solution containing either a protein or detergent blocking agent. After being probed with primary antibody, the membrane is washed and the antibody-antigen complexes are identified using horseradish peroxidase (HRPO) or alkaline phosphatase (AP) enzymes coupled to the secondary anti-immunoglobulin-G (anti-IgG) antibody (e.g., goat anti-rabbit IgG). The enzymes are attached directly (Basic Protocol) or via an avidin-biotin bridge (Alternate Protocol) to the secondary antibody. Chromogenic or luminescent substrates (Support Protocols 1 and 2) are then used to visualize the activity.

IMMUNOPROBING WITH DIRECTLY CONJUGATED SECONDARY ANTIBODY

BASIC PROTOCOL

After electrophoretic transfer to the membrane (*UNIT 10.7*), the immobilized proteins are probed with specific antibodies to identify and quantitate any antigens present. The membrane is first immersed in blocking buffer to fill all protein-binding sites with a nonreactive protein or detergent. Next, the membrane is placed in a solution containing an antibody directed against the antigen (primary antibody). The blot is washed and then exposed to an enzyme-antibody conjugate directed against the primary antibody (secondary antibody; e.g., goat anti-rabbit IgG). Antigens are identified by chromogenic or luminescent visualization (see Support Protocols 1 and 2) of the antigen/primary antibody/secondary antibody/enzyme complex bound to the membrane. Tween 20 is a common alternative to protein blocking agents for use with nitrocellulose or PVDF filters.

Materials

Membrane with transferred proteins (*UNIT 10.8*)
Blocking buffer (see recipe) appropriate for membrane and detection protocol
Primary antibody specific for protein of interest
TTBS (nitrocellulose or PVDF) *or* TBS (neutral or positively charged nylon; see recipes for both solutions)
Secondary antibody conjugate: horseradish peroxidase (HRPO)- or alkaline phosphatase (AP)-anti-Ig (Cappel, Vector, Kirkegaard & Perry, or Sigma; dilute as indicated by manufacturer and store frozen in 25-µl aliquots until use)
Heat-sealable plastic bags
Powder-free gloves
Plastic box

Copyright © 1996 by John Wiley & Sons, Inc.

BCU LIBRARY / CENTRAL UNIVERSITY LIBRARY

Additional reagents and equipment for chromogenic or luminescent visualization (see Support Protocol 1 or see Support Protocol 2)

1. Place membrane in heat-sealable plastic bag with 5 ml blocking buffer and seal bag. Incubate 30 min to 1 hr at room temperature with agitation on an orbital shaker or rocking platform.

   Usually 5 ml buffer is sufficient for two to three membranes (14 × 14–cm size).

   Plastic incubation trays are often used in place of heat-sealable bags, and can be especially useful when processing large numbers of strips in different primary antibody solutions.

2. Dilute primary antibody in blocking buffer.

   Primary antibody dilution is determined empirically but is typically 1/100 to 1/1000 for a polyclonal antibody (Fig. 10.10.1; Cooper and Paterson, 1995; Andrew and Titus, 1991a,b,c, 1993), 1/10 to 1/100 for hybridoma supernatants (Yokoyama, 1991a), and ≥1/1000 for murine ascites fluid containing monoclonal antibodies (Yokoyama, 1991b). Ten- to one-hundred-fold higher dilutions can be used with alkaline phosphatase– or luminescence–based detection systems. Both primary and secondary antibody solutions can be used at least twice, but long-term storage (i.e., >2 days at 4°C) is not recommended.

3. Open bag and pour out blocking buffer. Replace with diluted primary antibody and incubate 30 min to 1 hr at room temperature with constant agitation.

   Usually 5 ml diluted primary antibody solution is sufficient for two to three membranes (14 × 14–cm size). Incubation time may vary, depending on conjugate used.

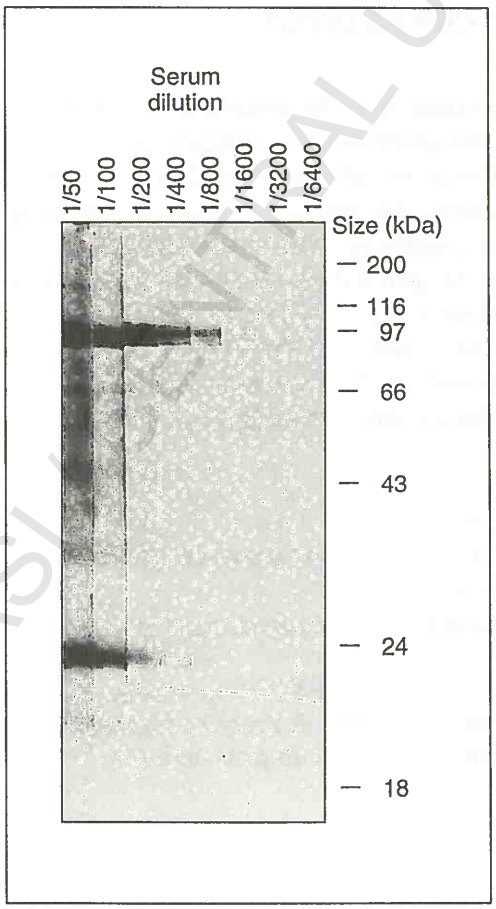


Figure 10.10.1 Serial dilution of primary antibody directed against the 97-kDa catalytic subunit of the plant plasma membrane ATPase. The blot was developed with HRPO-coupled avidin-biotin reagents according to the Alternate Protocol and visualized with 4-chloro-1-naphthol (4CN). Note how background improves with dilution.

BCU IASI CENTRAL UNIVERSITY LIBRARY

When using plastic trays, the primary and secondary antibody solution volume should be increased to 25 to 50 ml. For membrane strips, incubation trays with individual slots are recommended. Typically, 0.5 to 1 ml solution/slot is needed.

4. Remove membrane from plastic bag with gloved hand. Place in plastic box and wash 4 times by agitating with 200 ml TTBS (nitrocellulose or PVDF) or TBS (nylon), 10 to 15 min each time.

5. Dilute secondary antibody HRPO- or AP-anti-Ig conjugate in blocking buffer.

   Commercially available enzyme-conjugated secondary antibody is usually diluted 1/200 to 1/2000 (i.e., 20 µl/ml to 2 µl/ml) prior to use (Harlow and Lane, 1988).

6. Place membrane in fresh heat-sealable plastic bag, add diluted HRPO- or AP-anti-Ig conjugate, and incubate 30 min to 1 hr at room temperature with constant agitation.

   When using plastic incubation trays, see step 3 annotation for proper antibody solution volumes.

7. Remove membrane from bag and wash as in step 4. Develop according to appropriate visualization protocol (see Support Protocol 1 or see Support Protocol 2).

IMMUNOPROBING WITH AVIDIN-BIOTIN COUPLING TO SECONDARY ANTIBODY

ALTERNATE PROTOCOL

The following procedure is based on the Vectastain ABC kit from Vector (*SUPPLIERS APPENDIX*). It uses an avidin-biotin complex to attach horseradish peroxidase (HRPO) or alkaline phosphatase (AP) to the biotinylated secondary antibody. Avidin-biotin systems are capable of extremely high sensitivity because multiple reporter enzymes are bound to each secondary antibody. In addition, the detergent Tween 20 is a popular alternative to protein blocking agents when using nitrocellulose or PVDF membranes.

Additional Materials (also see Basic Protocol)

Vectastain ABC (HRPO) *or* ABC-AP (AP) kit (Vector) containing the following: reagent A (avidin), reagent B (biotinylated HRPO or AP), and biotinylated secondary antibody (request membrane immunodetection protocols when ordering)

1. Equilibrate membrane in appropriate blocking buffer in heat-sealed plastic bag with constant agitation using an orbital shaker or rocking platform. For nitrocellulose and PVDF, incubate 30 to 60 min at room temperature. For nylon, incubate ≥2 hr at 37°C.

   TTBS is well suited for avidin-biotin systems. For nylon, protein binding agents are recommended. Because nonfat dry milk contains residual biotin that will interfere with the immunoassay, its use must be restricted to the blocking step only.

   Plastic incubation trays are often used in place of heat-sealable bags, and can be especially useful when processing large numbers of strips in different primary antibody solutions.

2. Prepare primary antibody solution in TTBS (nitrocellulose or PVDF) or TBS (nylon).

   Dilutions of sera containing primary antibody generally range from 1/100 to 1/10,000. This depends in large part on the sensitivity of the detection system. With high-sensitivity avidin-biotin systems, dilutions from 1/1000 to 1/100,000 are common. Higher dilutions can be used with AP- or luminescence-based detection systems. To determine the appropriate concentration of primary antibody, a dilution series is easily performed with membrane strips. Separate antigens on a preparative gel (i.e., with a single large sample well) and immunoblot the entire gel. Cut 2- to 4-mm strips by hand or with a membrane cutter (Schleicher & Schuell or Inotech) and incubate individual

BOSTON UNIVERSITY LIBRARY

strips in a set of serial dilutions of primary antibody. The correct dilution should give low background and high specificity (Fig. 10.10.1).

3. Open bag, remove blocking buffer, and add enough primary antibody solution to cover membrane. Reseal bag and incubate 30 min at room temperature with gentle rocking.

   When using plastic trays, the primary and secondary antibody solution volume should be increased to 25 to 50 ml. For membrane strips, incubation trays with individual slots are recommended. Typically, 0.5 to 1 ml solution/slot is needed.

4. Remove membrane from bag and place in plastic box. Wash membrane 3 times over a 15-min span in TTBS (nitrocellulose or PVDF) or TBS (nylon). Add enough TTBS or TBS to fully cover the membrane (e.g., 5 to 10 ml/strip or 25 to 50 ml/whole membrane).

5. Prepare biotinylated secondary antibody solution by diluting 2 drops biotinylated antibody with 50 to 100 ml TTBS (nitrocellulose or PVDF) or TBS (nylon).

   This dilution gives both high sensitivity and enough volume to easily cover a large (14 × 14–cm) membrane.

6. Transfer membrane to fresh plastic bag containing secondary antibody solution. Incubate 30 min at room temperature with slow rocking, then wash as in step 4.

   When using plastic incubation trays, see step 3 annotation for proper antibody solution volumes.

7. While membrane is being incubated with secondary antibody, prepare avidin-biotin-HRPO or -AP complex. Mix 2 drops Vectastain reagent A and 2 drops reagent B into 10 ml TTBS (nitrocellulose or PVDF) or TBS (nylon). Incubate 30 min at room temperature, then further dilute to 50 ml with TTBS or TBS.

   Diluting the A and B reagents to 50 ml expands the amount of membrane that can be probed without greatly affecting sensitivity. Azide is a peroxidase inhibitor and should not be used as a preservative for long-term storage of the antibody solution. Casein, nonfat dry milk, serum, and some grades of bovine serum albumin (BSA) may interfere with the formation of the avidin-biotin complex and should not be used in the presence of avidin or biotin reagents (Gillespie and Hudspeth, 1991; see also instructions from Vector).

8. Transfer membrane to fresh plastic bag containing avidin-biotin-enzyme solution. Incubate 30 min at room temperature with slow rocking, then wash over a 30-min span as in step 4.

   Hybridization in a plastic bag requires 5 to 10 ml avidin-biotin-enzyme solution. Membrane strips require 5 to 10 ml/strip, whereas blots from standard-sized gels (i.e., 14 × 16 cm) require 50 ml for convenient handling in a tray.

9. Develop membrane according to the appropriate visualization protocol (see Support Protocol 1 or see Support Protocol 2).

BCU IAS/CENTRAL UNIVERSITY LIBRARY

VISUALIZATION WITH CHROMOGENIC SUBSTRATES

SUPPORT PROTOCOL 1

After incubation with primary and secondary antibody conjugates (see Basic Protocol or see Alternate Protocol), bound antigens are typically visualized with chromogenic substrates. The substrates 4CN, DAB/$NiCl_2$, and TMB are commonly used with horseradish peroxidase (HRPO)–based immunodetection procedures, whereas BCIP/NBT is recommended for alkaline phosphatase (AP)–based procedures (see Table 10.10.1). After incubation with primary and secondary antibodies, the membrane is placed in the appropriate substrate solution. Protein bands usually appear within a few minutes.

Table 10.10.1 Chromogenic and Luminescent Visualization Systems[a]

System	Reagent[b]	Reaction/detection	Comments[c]
Chromogenic			
HRPO-based	4CN	Oxidized products form purple precipitate	Not very sensitive (Tween 20 inhibits reaction); fades rapidly on exposure to light
	DAB/$NiCl_2$[d]	Forms dark brown precipitate	More sensitive than 4CN but potentially carcinogenic; resulting membrane is easily scanned
	TMB[e]	Forms dark purple stain	More stable and less toxic than DAB/$NiCl_2$; may be somewhat more sensitive[e]; can be used with all membrane types
AP-based	BCIP/NBT	BCIP hydrolysis produces indigo precipitate after oxidation with NBT; reduced NBT precipitates; dark blue-gray stain results	More sensitive and reliable than other AP-precipitating substrates; note that phosphate inhibits AP activity
Luminescent			
HRPO-based	Luminol/H_2O_2/ *p*-iodophenol	Oxidized luminol substrate gives off blue light; *p*-iodophenol increases light output	Very convenient, sensitive system; reaction is detected within a few seconds to 1 hr
AP-based	Substituted 1,2-dioxetane phosphates (e.g., AMPPD, CSPD, Lumigen-PPD, Lumi-Phos 530[f])	Dephosphorylated substrate gives off light	Protocol described gives reasonable sensitivity on all membrane types; consult instructions of reagent manufacturer for maximum sensitivity and minimum background (see Troubleshooting)

[a]Abbreviations: AMPPD or Lumigen-PPD, disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2′-tricyclo[3.3.1.1^{3,7}]-decan}-4-yl)phenyl phosphate; AP, alkaline phosphatase; BCIP, 5-bromo-4-chloro-3-indolyl phosphate; 4CN, 4-chloro-1-napthol; CSPD, AMPPD with substituted chlorine moiety on adamantine ring; DAB, 3,3′-diaminobenzidine; HRPO, horseradish peroxidase; NBT, nitroblue tetrazolium; TMB, 3,3′,5,5′-tetramethylbenzidine.

[b]Recipes and suppliers for all reagents except TMB are listed in Reagents and Solutions. Kits containing TMB are available from Kirkegaard & Perry, TSI Center for Diagnostic Products, Moss, and Vector.

[c]See Commentary for further details.

[d]DAB/$NiCl_2$ can be used without the nickel enhancement, but sensitivity is greatly reduced.

[e]First treating nitrocellulose filters with 1% dextran sulfate for 10 min in 10 mM citrate-EDTA (pH 5.0) causes TMB to precipitate onto the membrane with a sensitivity much greater than that seen for 4CN or DAB and equal to or better than that for BCIP/NBT (McKimm-Breschkin, 1990).

[f]Lumi-Phos 530 contains dioxetane, $MgCl_2$, cetyltrimethylammonium bromide (CTAB), and fluorescent enhancer in a pH 9.6 buffer.

BCU ASTON UNIVERSITY LIBRARY

Materials

Membrane with transferred proteins and probed with antibody-enzyme complex (see Basic Protocol or see Alternate Protocol)
TBS (see recipe)
Chromogenic visualization solution (Table 10.10.1; see recipes)

Additional reagents and equipment for gel photography (*UNIT 10.5*)

1. If final membrane wash (see Basic Protocol, step 7, or see Alternate Protocol, step 9) was performed in TTBS, wash membrane 15 min at room temperature in 50 ml TBS.

   The Tween 20 in the TTBS interferes with 4CN development (Bjerrum et al., 1988).

2. Place membrane into chromogenic visualization solution. Bands should appear in 10 to 30 min.

3. Terminate reaction by washing membrane in distilled water. Air dry and photograph (*UNIT 10.5*) for a permanent record.

SUPPORT PROTOCOL 2

VISUALIZATION WITH LUMINESCENT SUBSTRATES

After incubation with primary and secondary antibody conjugates (see Basic Protocol and see Alternate Protocol), antigens can also be visualized with luminescent substrates. Detection with light offers both greater speed and enhanced sensitivity over chromogenic and radioisotopic procedures. After the final wash, the blot is immersed in a substrate solution containing luminol for horseradish peroxidase (HRPO) systems, or dioxetane phosphate for alkaline phosphatase (AP) systems, sealed in thin plastic wrap and placed firmly against film. Exposures range from a few seconds to several hours, although typically strong signals appear within a few seconds or minutes.

Additional Materials (also see Support Protocol 1)

Luminescent substrate buffer: 50 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*; HRPO), *or* dioxetane phosphate substrate buffer (see recipe; AP)
Nitro-Block solution (AP reactions only): 5% (v/v) Nitro-Block (Tropix) in dioxetane phosphate substrate buffer (see recipe), prepared just before use
Luminescent visualization solution (Table 10.10.1; see recipes)
Clear plastic wrap

NOTE: See Troubleshooting section for suggestions concerning optimization of the protocol, particularly when employing AP-based systems.

1. Equilibrate membrane in two 15-min washes with substrate buffer each.

   For blots of whole gels, use 50 ml substrate buffer; for strips, use 5 to 10 ml/strip.

2. *For AP reactions using nitrocellulose or PVDF membranes:* Incubate 5 min in Nitro-Block solution, followed by 5 min in substrate buffer.

   For blots of whole gels, use 50 ml Nitro-Block solution and substrate buffer; for strips, use 5 to 10 ml/strip.

   Nitro-Block enhances light output from the dioxetane substrate in reactions using AMPPD, CSPD, or Lumigen-PPD concentrate. It is required for nitrocellulose and recommended for PVDF membranes. It is not needed for Lumi-Phos 530, AP reactions on nylon membranes, or HRPO–based reactions on any type of membrane. Lumi-Phos 530 is not recommended for nitrocellulose membranes.

3. Transfer membrane to luminescent visualization solution. Soak 30 sec (HRPO reactions) to 5 min (AP reactions).

BCU IASI CENTRAL UNIVERSITY LIBRARY

Alternatively, lay out a square of plastic wrap and pipet 1 to 2 ml visualization solution into the middle. Place membrane on the plastic so that the visualization solution spreads out evenly from edge to edge. Fold wrap back onto membrane, seal, and proceed to step 5.

4. Remove membrane, drain, and place face down on a sheet of clear plastic wrap. Fold wrap back onto membrane and seal with tape to form a liquid-tight enclosure.

   To ensure an optimal image, only one layer of plastic should be present between the membrane and film. Sealable bags are an effective alternative. Moisture must not come in contact with the X-ray film.

5. In a darkroom, place membrane face down onto film.

   Do this quickly and do not reposition; a double image will be formed if the membrane is moved while in contact with the film. A blurred image is usually caused by poor contact between membrane and film; use a film cassette that ensures a tight fit.

6. Expose film for a few seconds to several hours.

   Typically, immunoblots produce very strong signals within a few seconds or minutes; however, weak signals may require several hours to an overnight exposure. If no image is detected, expose film 30 min to 1 hr and, if needed, overnight (see Troubleshooting).

7. If desired, wash membrane in two 15-min washes of 50 ml TBS and process for chromogenic development (see Support Protocol 1).

   It is possible to develop the same membrane with chromogenic substrates after luminescent visualization.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Alkaline phosphate substrate buffer

100 mM Tris·Cl, pH 9.5
100 mM NaCl
5 mM $MgCl_2$

Blocking buffer

Colorimetric detection

For nitrocellulose and PVDF: 0.1% (v/v) Tween 20 in TBS (TTBS; see recipe).

For neutral and positively charged nylon: Tris-buffered saline (TBS; see recipe) containing 10% (w/v) nonfat dry milk.

TTBS can be stored up to 1 week at 4°C. Prepare blocking buffer containing nonfat dry milk immediately prior to use, as the milk blocking solution is not stable.

Luminescent detection

For nitrocellulose, PVDF, and neutral nylon (e.g., Pall Biodyne A): 0.2% (w/v) casein (e.g., Hammarsten grade or I-Block; Tropix) in TTBS (see recipe).

For positively charged nylon: 6% (w/v) casein/1% (v/v) polyvinylpyrrolidone (PVP) in TTBS (see recipe).

For each solution: With constant mixing, add casein and PVP to warm (65°C) TTBS. Stir for 5 min, then cool. Prepare each solution just before use.

BOYLAN/CENTRAL UNIVERSITY LIBRARY

Chromogenic visualization solutions

BCIP/NBT visualization solution: Mix 33 µl NBT stock [100 mg NBT in 2 ml 70% (v/v) dimethylformamide (DMF), stored <1 year at 4°C] and 5 ml alkaline phosphate substrate buffer (see recipe). Add 17 µl BCIP stock (100 mg BCIP in 2 ml 100% DMF, stored <1 year at 4°C) and mix. Stable 1 hr at room temperature.

Recipe is from Harlow and Lane (1988). Alternatively, BCIP/NBT substrates may be purchased from Sigma, Kirkegaard & Perry, Moss, and Vector (SUPPLIERS APPENDIX).

4CN visualization solution: Mix 20 ml ice-cold methanol with 60 mg 4-chloro-1-naphthol (4CN). Separately mix 60 µl of 30% (w/v) H_2O_2 with 100 ml TBS (see recipe) at room temperature. Rapidly mix the two solutions and use immediately.

DAB/$NiCl_2$ visualization solution:
5 ml 100 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*)
100 µl DAB stock (40 mg/ml in H_2O, stored in 100-µl aliquots at –20°C)
25 µl $NiCl_2$ stock (80 mg/ml in H_2O, stored in 100-µl aliquots at –20°C)
15 µl 3% (w/v) H_2O_2
Mix just before use

CAUTION: *Handle DAB carefully, wearing gloves and mask; it is a carcinogen.*

Suppliers of chromogenic HRPO substrates (4CN and DAB/$NiCl_2$) are Sigma, Kirkegaard & Perry, Moss, and Vector (SUPPLIERS APPENDIX). For selection of appropriate chromogenic solutions, and for definition of abbreviations, see Table 10.10.1.

Dioxetane phosphate substrate buffer

1 mM $MgCl_2$
0.1 M diethanolamine
0.02% (w/v) sodium azide (optional)
Adjust to pH 10 with HCl and use fresh

Traditionally, the AMPPD substrate buffer has been a solution containing 1 mM $MgCl_2$ and 50 mM sodium carbonate/bicarbonate, pH 9.6 (Gillespie and Hudspeth, 1991). The use of diethanolamine results in better light output (Tropix Western Light instructions).

Alternatively, 100 mM Tris·Cl (pH 9.5)/100 mM NaCl/5 mM $MgCl_2$ can be used (Sandhu et al., 1991).

Luminescent visualization solutions

Dioxetane phosphate visualization solution: Prepare 0.1 mg/ml AMPPD or CSPD (Tropix) *or* 0.1 mg/ml Lumigen-PPD (Lumigen) substrate in dioxetane phosphate substrate buffer (see recipe). Prepare just before use. Lumi-Phos 530 (Boehringer Mannheim or Lumigen) is a ready-to-use solution and can be applied directly to the membrane.

This concentration (240 µM) of AMPPD substrate is the minimum recommended by Tropix. Ten-fold lower concentrations can be used but require longer exposures.

Luminol visualization solution:
0.5 ml 10× luminol stock [40 mg luminol (Sigma) in 10 ml dimethyl sulfoxide (DMSO)]
0.5 ml 10× *p*-iodophenol stock [optional; 10 mg (Aldrich) in 10 ml DMSO]
2.5 ml 100 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*)
25 µl 3% (w/v) H_2O_2
H_2O to 5 ml
Prepare just before use

Recipe is from Schneppenheim et al. (1991). Premixed luminol substrate mix (Mast Immunosystems, Amersham, Du Pont NEN Renaissance, or Kirkegaard & Perry Lumi-

continued

BCU IASI / CENTRAL UNIVERSITY LIBRARY

GLO) may also be used. For selection of appropriate luminescent solutions, and for definition of abbreviations, see Table 10.10.1.

p-Iodophenol is an optional enhancing agent that increases light output. Luminol and p-iodophenol stocks can be stored ≤6 months at –20°C.

Tris-buffered saline (TBS)

100 mM Tris·Cl, pH 7.5 (*APPENDIX 2E*)
0.9% (w/v) NaCl
Store up to several months at 4°C

Tween 20/TBS (TTBS)

0.1% (v/v) Tween 20 in Tris-buffered saline (TBS; see recipe)
Store up to several months at 4°C

COMMENTARY

Background Information

Immunoprecipitation has been widely used to visualize the antigens recognized by various antibodies, both polyclonal and monoclonal. However, there are several problems inherent to immunoprecipitation, including the need to radiolabel the antigen, coprecipitation of tightly associated macromolecules, occasional difficulty in obtaining precipitating antibodies, and insolubility of various antigens (Talbot et al., 1984).

To circumvent these problems, electroblotting (Towbin et al., 1979)—subsequently popularized as western blotting or immunoblotting (Burnette, 1981)—was conceived. Immunoblotting is a rapid and sensitive assay for the detection and characterization of proteins that works by exploiting the specificity inherent in antigen-antibody recognition. It involves the solubilization and electrophoretic separation of proteins, glycoproteins, or lipopolysaccharides by SDS-PAGE (*UNITS 10.1 & 10.4*) or urea-PAGE (*UNIT 10.2*), followed by quantitative transfer and irreversible binding to nitrocellulose, PVDF, or nylon membranes (*UNIT 10.8*). This technique has been useful in identifying specific antigens recognized by polyclonal or monoclonal antibodies, and is highly sensitive (1 ng of antigen can be detected).

Protein blotting has important clinical applications—notably, it is the confirmatory test for human immunodeficiency virus type 1 (HIV-1). SDS-PAGE-separated virus proteins are blotted onto nitrocellulose and processed with patient sera. After washing and incubating (using a Tween 20/nonfat dry milk diluting and washing solution) with an anti-human IgG coupled to horseradish peroxidase (HRPO) or alkaline phosphatase (AP), the antigens are identified by chromogenic development. Typically, the prepared blots and all reagents needed for the test are purchased commercially (Bio-Rad Novapath, Organon Teknika Cappel HIV-1 Western Blot Kit, or Du Pont NEN HIV Western Blot Kit).

Immunoblotted proteins can be detected by chromogenic or luminescent assays (see Table 10.10.1 for a description of the reagents available for each system, their reactions, and a comparison of their advantages and disadvantages). Luminescent detection methods offer several advantages over traditional chromogenic procedures. In general, luminescent substrates increase the sensitivity of both HRPO and AP systems without the need for radioisotopes. Substrates for the latter have only recently been applied to protein blotting (see Gillespie and Hudspeth, 1991; Sandhu et al., 1991; Bronstein et al., 1992). Luminescent detection can be completed in as little as a few seconds; exposures rarely exceed 1 hr. Depending on the system, the luminescence can last up to 3 days, permitting multiple exposures of the same blot. Furthermore, the signal is detected by film, and varying the exposure can result in more or less sensitivity. Luminescent blots can be easily erased and reprobed because the reaction products are soluble and do not deposit on the membrane. Compared to chromogenic development, a luminescent image recorded on film is easier to photograph and to quantitate by densitometry.

AP–based luminescent protocols that achieve maximum sensitivity with minimum background can be complex, and the manufacturer's instructions should be consulted (see Reagents and Solutions). The procedure described in Support Protocol 2 gives reasonable sensitivity on nitrocellulose, PVDF, and nylon membranes with a minimum of steps.

B. C. U. „M. EMINESCU" IAŞI
CHIMIE

BOSTON UNIVERSITY LIBRARY

Critical Parameters

First and foremost, the antibody being used should recognize denatured antigen. Nonspecific binding of antibodies can occur, so control antigens and antibodies should always be run in parallel. Time of transfer and dilutions of primary antibody and conjugate should always be optimized.

A variety of agents are currently used to block binding sites on the membrane after blotting (Harlow and Lane, 1988). These include Tween 20, PVP, nonfat dry milk, casein, BSA, and serum. A 0.1% (v/v) solution of Tween 20 in TBS (TTBS), a convenient alternative to protein-based blocking agents, is recommended for chromogenic development of nitrocellulose and PVDF membranes (Blake et al., 1984). In contrast to dry milk/TBS blocking solution, TTBS is stable and has a long shelf life at 4°C. Furthermore, TTBS generally produces a clean background and permits subsequent staining with India ink.

Two types of nylon membranes are used for transfer—neutral (e.g., Pall Biodyne A) and positively charged (e.g., Pall Biodyne B). Although the positively charged membranes have very good protein-binding characteristics, they tend to have a higher background. These membranes remain positively charged from pH 3 to pH 10. Neutral nylon membranes are also charged, having a mix of amino and carboxyl groups that give an isoelectric point of 6.5. Because of their high binding capacity, positively charged membranes are popular for protein applications using luminescence.

Nylon membranes require more stringent blocking steps. Here 10% (w/v) nonfat dry milk in TBS is recommended for chromogenic development. During development of luminescence, however, background is a more significant problem. When compared to dry milk, purified casein has minimal endogenous AP activity (such activity leads to high background) and is therefore recommended as a blocking agent for nitrocellulose, PVDF, and nylon membranes. Positively charged nylon requires much more stringent blocking with 6% (w/v) casein and 1% (v/v) polyvinylpyrrolidone (PVP-40). Because nonfat dry milk and casein may contain biotin, which will interfere with avidin-biotin reactions, subsequent steps are done without protein blocking agents when using these systems. If background is a problem, highly purified casein (0.2% to 6%) added to the antibody incubation buffers may help.

Troubleshooting

There are several problems associated with immunoblotting. The antigen is solubilized and electrophoresed in the presence of denaturing agents (e.g., SDS or urea), and some antibodies may not recognize the denatured form of the antigen transferred to the membrane. The results observed may be entirely dependent on the denaturation and transfer system used. For example, zwitterionic detergents have been shown to restore the antigenicity of outer membrane proteins in immunoblotting (Mandrell and Zollinger, 1984). Gel electrophoresis under nondenaturing conditions can sometimes preserve antigenicity (*UNIT 10.3*).

Other potential problems include high background, nonspecific or weak cross-reactivity of antibodies, poor protein transfer or membrane binding efficiency, and insufficient sensitivity. For an extensive survey and discussion of immunoblotting problems and artifacts, see Bjerrum et al. (1988).

Insufficient blocking or nonspecific binding of the primary or secondary antibody will cause a high background stain. A control using preimmune sera or only the secondary antibody will determine if these problems are due to the primary antibody. Try switching to another blocking agent; protein blocking agents may weakly cross-react. Lowering the concentration of primary antibody should decrease background and improve specificity (Fig. 10.10.1).

Because of the nature of light and the method of detection, certain precautions are warranted when using luminescent visualization (e.g., Harper and Murphy, 1991). Very strong signals can overshadow weaker signals nearby on the membrane. Because light will pipe through the membrane and the surrounding plastic wrap, overexposure will produce a broad, diffuse image on the film. The signal can also saturate the film, exposing it to a point whereby increased exposure will not cause a linear increase in the density of the image on the film.

With the AP substrate AMPPD, nitrocellulose, PVDF, and nylon membranes require 2, 4, and 8 to 12 hr, respectively, to reach maximum light emission. In addition, PVDF is reported to give a stronger signal than nitrocellulose (Tropix Western Light instructions). Positively charged nylon requires special blocking procedures to minimize background (Gillespie and Hudspeth, 1991). These procedures include using a blocking and pri-

RCU IAS CENTRAL LIBRARY

mary antibody solution containing 6% (w/v) casein, 1% (v/v) polyvinylpyrrolidone-40 (PVP-40), 3 mM NaN_3, 10 mM EDTA, and phosphate-buffered saline (PBS), pH 6.8. Prior to use, the casein must be heated to 65°C to reduce AP activity in the casein itself. In addition, maximum sensitivity has been observed when free biotin or biotinylated proteins are removed by pretreating the casein with avidin-agarose (Sigma).

Anticipated Results

Immunoblotting should result in the detection of one or more bands. Although antibodies directed against a single protein should produce a single band, degradation of the sample (e.g., via endogenous proteolytic activity) may cause visualization of multiple bands of slightly different size. Multimers will also form spontaneously, causing higher-molecular-weight bands on the blot. If one is simultaneously testing multiple antibodies directed against a complex protein mixture (e.g., using a patient sera against SDS-PAGE-separated viral proteins in the AIDS western blot test), multiple bands will be visualized.

Time Considerations

The entire immunoblotting procedure can be completed in 1 to 2 days, depending on the transfer time and type of gel. Gel electrophoresis requires 4 to 6 hr on a regular gel and 1 hr on a minigel. Transfer time can be 1 hr (for high-power transfer) to overnight. Blocking, conjugate incubation, and washing each take 30 min to 1 hr. Finally, substrate incubation requires 10 to 30 min (chromogen) and a few seconds to several hours (luminescence).

Literature Cited

Andrew, S.A. and Titus, J.A. 1991a. Purification of immunoglobulin G. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.7.1-2.7.12. John Wiley & Sons, New York.

Andrew, S.A. and Titus, J.A. 1991b. Fragmentation of immunoglobulin G. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.8.1-2.8.10. John Wiley & Sons, New York.

Andrew, S.A. and Titus, J.A. 1991c. Fragmentation of immunoglobulin M. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.10.1-2.10.4. John Wiley & Sons, New York.

Andrew, S.A. and Titus, J.A. 1993. Purification of immunoglobulin M. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.9.1-2.9.3. John Wiley & Sons, New York.

Bjerrum, O.J., Larsen, K.P., and Heegaard, N.H.H. 1988. Nonspecific binding and artifacts—specificity problems and troubleshooting with an atlas of immunoblotting artifacts. *In* CRC Handbook of Immunoblotting of Proteins, Vol. I: Technical Descriptions (O.J. Bjerrum and N.H.H. Heegaard, eds.) pp. 227-254. CRC Press, Boca Raton, Fla.

Blake, M.S., Johnston, K.H., Russell-Jones, G.J., and Gotschlich, E.C. 1984. A rapid, sensitive method for detection of alkaline phosphatase–conjugated anti-antibody on Western blots. *Anal. Biochem.* 136:175-179.

Bronstein, I., Voyta, J.C., Murphy, O.J., Bresnick, L., and Kricka, L.J. 1992. Improved chemiluminescent Western blotting procedure. *BioTechniques* 12:748-753.

Burnette, W.N. 1981. Western blotting: Electrophoretic transfer of proteins from sodium dodecyl sulfate–polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112:195-203.

Cooper, H.M. and Paterson, Y. 1995. Production of polyclonal antisera. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.4.1-2.4.9. John Wiley & Sons, New York.

Gillespie, P.G. and Hudspeth, A.J. 1991. Chemiluminescence detection of proteins from single cells. *Proc. Natl. Acad. Sci. U.S.A.* 88:2563-2567.

Harlow, E. and Lane, D. 1988. Immunoblotting. *In* Antibodies: A Laboratory Manual, pp. 471-510. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Harper, D.R. and Murphy, G. 1991. Nonuniform variation in band pattern with luminol/horseradish peroxidase Western blotting. *Anal. Biochem.* 192:59-63.

Mandrell, R.E. and Zollinger, W.D. 1984. Use of zwitterionic detergent for the restoration of antibody-binding capacity of electroblotted meningococcal outer membrane proteins. *J. Immunol. Methods* 67:1-11.

McKimm-Breschkin, J.L. 1990. The use of tetramethylbenzidine for solid phase immunoassays. *J. Immunol. Methods* 135:277-280.

Sandhu, G.S., Eckloff, B.W., and Kline, B.C. 1991. Chemiluminescent substrates increase sensitivity of antigen detection in Western blots. *BioTechniques* 11:14-16.

Schneppenheim, R., Budde, U., Dahlmann, N., and Rautenberg, P. 1991. Luminography—a new, highly sensitive visualization method for electrophoresis. *Electrophoresis* 12:367-372.

Talbot, P.V., Knobler, R.L., and Buchmeier, M. 1984. Western and dot immunoblotting analysis of viral antigens and antibodies: Application to

BOSTON UNIVERSITY MEDICAL CENTER HOSPITAL LIBRARY

murine hepatitis virus. *J. Immunol. Methods* 73:177-188.

Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Yokoyama, W. 1991a. Production of monoclonal antibodies. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.5.1-2.5.17. John Wiley & Sons, New York.

Yokoyama, W. 1991b. Monoclonal antibody supernatant and ascites fluid. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.6.1-2.6.7. John Wiley & Sons, New York.

Key References

Gillespie and Hudspeth, 1991. See above.

Describes alkaline phosphatase–based luminescent detection methods.

Harlow and Lane, 1988. See above.

Details alternative detection methods.

Schneppenheim et al., 1991. See above.

Details horseradish peroxidase–based luminescent detection methods.

Contributed by Sean Gallagher
Hoefer Scientific Instruments
San Francisco, California

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Autoradiography

Autoradiography is a radioisotope-detection method that records the spatial distribution of radioisotope-labeled substances within a membrane or specimen material. The ionizing radiation (usually β particles) that is emitted from the radioisotope interacts with the silver halide grains within the emulsion of the photographic film, forming a latent image. The latent image is amplified by the action of the developer during processing, thus creating a visible image. The pattern and density of the image are used to locate and quantify the radioactive distribution.

Commercially available autoradiography films are best suited for autoradiography because of their high sensitivity to both light and ionizing radiation. Autoradiography films generally yield higher-contrast images with a shorter exposure time than medical X-ray films and other commercially available photographic films.

The most appropriate choice of autoradiography film depends upon the detection requirements of the experiment and the type and quantity of the radiolabeled isotope. Several autoradiography films are designed and used with very specific exposure methods and radioisotopes—e.g., BioMax MR and BioMax MS (Eastman Kodak). BioMax MR is designed for direct-exposure detection of ^{14}C and ^{35}S radioisotopes and BioMax MS is designed for exposure detection of ^{32}P and ^{125}I radioisotopes with an intensifying screen. The following descriptions of films and exposure procedures will assist in optimizing autoradiography techniques.

TYPES OF FILM

The structure of autoradiography film includes either one emulsion layer or two on a 7-mil (178-μm) polyester support. Double-coated film includes emulsion layers on both sides of the sheet of film. Emulsion layers consist of light-sensitive silver halide grains and gelatin. The polyester support can be either clear or blue-tinted.

Single-Coated Films

Autoradiography films that are single-emulsion coated are optimized for use with medium-energy radioisotopes—e.g., ^{14}C, ^{35}S, and ^{33}P—and direct-exposure techniques (see discussion of Direct Exposure) for a wide range of biological applications such as DNA sequencing. At maximum energy levels of 150 to 170 KeV, the majority of the β particles do not penetrate through the film base, so an additional emulsion on the other side of the film support, such as that found in double-coated films, does not increase sensitivity to ^{14}C and ^{35}S radioisotopes. Single-coated films—e.g., BioMax MR—have greater clarity because they do not have the second emulsion layer. BioMax MR is highly sensitive to direct radioisotope exposure. It is twice as sensitive to ^{14}C and ^{35}S as X-Omat AR film (Eastman Kodak; see discussion of Double-Coated Films). BioMax MR film sensitivity to light has been reduced in order to increase allowable safelight handling time in the darkroom, thus reducing its sensitivity to fluorography or intensifying-screen exposures.

Double-Coated Films

Double-coated films are generally used with intensifying screens (see discussion of Intensifying-Screen Exposure) at reduced temperatures to detect high-energy radioisotopes such as ^{32}P and ^{125}I. High-energy β particles, X rays, and γ rays penetrate and expose both emulsion layers. These high-energy particles are primarily absorbed by an intensifying screen if it is placed on the opposite side of the film from the specimen. Light is emitted by the phosphor grains in the intensifying screen upon absorbing the β particle, and this greatly amplifies detection. Double-coated X-Omat AR film is a widely used film for general autoradiography. This film is highly sensitive to blue light and is used with a blue-light-emitting calcium tungstate ($CaWO_4$) intensifying screen.

The green-light-sensitive BioMax MS film is a double-coated film that gives greatest sensitivity to ^{32}P when used with a BioMax MS intensifying screen or high-energy TranScreen (Eastman Kodak) at –70°C. Both of these screens emit green light when exposed to ionizing radiation. These film/screen systems are approximately four times more sensitive than the X-Omat AR film/$CaWO_4$ intensifying screen system.

Most fluorographic exposures (see discussion of Fluorographic Exposure) emit light in the blue spectrum. X-Omat AR film is a blue-sensitive film that is generally used with scintillants such as PPO (2, 5-diphenyloxazole) and

Contributed by Daniel C. Bundy
Current Protocols in Protein Science (1997) 10.11.1-10.11.6
Copyright © 1997 by John Wiley & Sons, Inc.

BIOMEDICAL SCIENCES LIBRARY

sodium salicylate, as well as prepared fluorographic solutions and sprays (Amersham, DuPont NEN). X-Omat AR film is at least two times more sensitive than other blue-sensitive autoradiography films for most fluorographic exposure techniques.

HANDLING OF AUTORADIOGRAPHY FILM

Safelighting

Safe darkroom illumination is dependent upon the type of safelight filter, the wattage of the bulb, and the distance from the safelight bulb to the sheet of film. A red safelight filter (e.g., Kodak GBX-2) is recommended for the majority of autoradiography films. The bulb wattage should not exceed 15 W with direct-illuminating safelight lamps. The minimum safelight-to-film distance is 4 feet.

If one is unsure of the safelight conditions in a darkroom or the safelight sensitivity of a film, it is wise to perform a safelight safety test. In the dark, put the film on the workbench and place a coin on top of the film. Next, turn on the safelight for 1 to 2 min or more, then turn off the safelight and process the film in the dark. If the coin's shadow image is not visible on the film, then the safelight conditions are suitable for the tested exposure time.

If the safelight safety test fails, measures to be taken may include replacing the safelight filter, decreasing the bulb wattage, increasing the safelight-to-film distance, or decreasing the time of exposure to safelight conditions.

Storage and Care of Film

BioMax MS, BioMax MR, and X-Omat AR films are highly efficient detectors of background radiation. With these films, the background fog can be expected to increase with film age. It is advisable to keep down the inventory of these films and order them more frequently to minimize background fog. Most laboratories routinely order just enough film for 3 to 5 months.

Film sheets should be handled carefully to prevent black or white marks and lines that result from pinching and scratching the emulsion. Individual film sheets should be handled only by the edges.

QUANTITATIVE AUTORADIOGRAPHY

A radioactive calibration strip is very effective in autoradiography quantification techniques. Commercially available radioactive standards (e.g., ^{3}H and ^{14}C tissue equivalent standards from American Radiolabeled Chemicals or Amersham), self-prepared radioactive scales, or radioisotope-labeled dilution series are used during an autoradiographic exposure to expose the film or storage phosphor screen to quantitate the image. The autoradiography image (e.g., from a tissue specimen, membrane, or gel) and the calibration strip require digitization or densitometry to convert the information to radioactivity from a plot of radioactivity versus density or pixel value.

A common method of preparing a calibration strip is to serially dilute the radioisotope label by a factor of 2 or more and apply it to a substrate or membrane. It is important to extend the dilution series beyond the estimated range of the radioisotope label concentration. The individual radioisotope-label concentrations are spotted (3 to 5 µl), blotted, or hybridized onto the support material. The shape and configuration of each spot or band should be similar and uniform. The density or pixel value is plotted versus the radioactivity value to obtain the relative radioisotope label calibration curve.

EXPOSURE PROCEDURES

There are many factors to consider when choosing an autoradiographic exposure procedure. The choices include the isotope, the radioactivity level, whether to use double- or single-coated film, film treatment, and exposure temperature. Each researcher needs to determine the best procedure for each particular experimental protocol (see Table 10.11.1).

Direct Exposure

For direct exposure, the radiolabeled specimen is placed in direct contact with autoradiography film and enclosed in a light-tight cassette for a predetermined, or suitable, amount of time. At the end of exposure, the film is removed from the cassette after noting the orientation of the film sheet with respect to the specimen by bending a corner of the film towards the specimen or by marking a corner of the film with a no. 2 pencil. The film is then photographically processed.

Exposing autoradiography film at reduced temperatures (see discussion of Fluorographic Exposure) does not improve the sensitivity with direct β-particle exposures as it does with fluorographic and intensifying-screen exposures. The β particle forms a very stable latent image in the silver grain of the emulsion. This latent image does not significantly fade with time as in the case with light exposures from fluors

BCU IAS (CENTRAL LIBRARY)

Table 10.11.1 Autoradiography Exposure Procedures

Application(s)	Isotope	Priority	Exposure procedure[a]
Protein gels, western blots	^{3}H, ^{14}C, ^{35}S	Speed	Fluorography TranScreen LE
Nucleic acid sequencing	^{32}P	Resolution Speed	Direct
	^{35}S	Resolution Speed	Direct Fluorography TranScreen LE
Oligonucleotide separation	^{32}P	Resolution Speed	Direct Intensifying screen
Northern, Southern, and dot blots; library screening	^{32}P ^{35}S	Speed Speed	Intensifying screen Fluorography TranScreen LE
In situ hybridization	^{14}C, ^{35}S	Resolution Speed	Direct Fluorography TranScreen LE
	^{3}H	Speed	Fluorography TranScreen LE

[a]TranScreen LE is a product of Eastman Kodak (see *SUPPLIERS APPENDIX*).

(scintillants) and intensifying screens (see discussions of Intensifying-Screen Exposure and Fluorographic Exposure).

Direct-exposure techniques are effective with all β-particle-emitting radioisotopes except tritium (^{3}H). Tritium does not have enough energy to penetrate the overcoat layer on top of autoradiography film and expose the silver halide emulsion. ^{14}C and ^{35}S direct-exposure techniques provide the highest imaging resolution when used with a single-coated autoradiographic film such as BioMax MR film (Steinfeld et al., 1994), although longer exposure times are required than for fluorographic exposures with X-Omat AR film.

Wet or moist specimens should be sealed in a plastic bag or wrapped in plastic wrap. Water and other solvents should not come in contact with film. The plastic wrap attenuates radiation from medium-energy β emitters (^{35}S or ^{14}C) by a factor of 2. The effect with ^{32}P is less, but still noticeable. The investigator should be aware of the thickness of the plastic wrap used. Commercially available plastic wraps are acceptable; extremely thin plastic wraps are better but more expensive.

Fluorographic Exposure

Fluorography enhances the detection of low levels of low- and medium-energy radioisotopes such as ^{3}H, ^{14}C, and ^{35}S. The scintillant (i.e., a reagent that emits visible light upon absorbing ionizing radiation) is placed either on or in the radiolabeled specimen. When the scintillant absorbs the β-particle energy it emits blue light, which is detected by autoradiographic film.

In fluorographic exposure, the radiolabeled specimen (e.g., gel, nitrocellulose membrane, or TLC plate) is soaked or sprayed with the scintillant. The specimen is then dried according to the instructions provided by the scintillant manufacturer and published protocols. Next, under safelight conditions, the specimen is placed against autoradiography film in a light-tight cassette. The temperature for optimum exposure sensitivity is between –70° and –100°C.

At the end of the exposure, the cassette is warmed to ambient temperature before opening. The film is carefully removed from the specimen after marking the orientation of the film sheet with respect to the specimen by bending a corner of the film towards the specimen or by marking a corner of the film with a no. 2 pencil. Then the film is photographically processed. If the film and specimen are separated too rapidly, static discharge may occur. This static may appear on the film as many small black dots or as large, branched black streaks.

X-Omat AR film has a spectral sensitivity that is optimal for detecting fluorescence from

BUFFALO GENERAL UNIVERSITY LIBRARY

the scintillant 2,5-diphenyloxazole (PPO), which emits at 388 nm, sodium salicylate, which emits at 420 nm, and other commercially available fluorographic solutions and sprays. Commercially available scintillants and fluorographic products are available from several companies, including Amersham Corporation and DuPont NEN.

The invisible latent image is the forerunner of the visible, metallic silver image that arises from photographic development after exposure to ionizing radiation. This darkening of the photographic film is the result of the latent image, on the silver halide grain, being converted to metallic silver by reducing agents in the developer. Two or more photons of light are required to make a latent image center on a silver halide grain, whereas only one medium-energy β particle (from ^{14}C or ^{35}S) usually forms many latent image centers on numerous silver halide grains. Generally the latent image centers formed from the interaction with two or more light photons are not stable and fade via low-intensity reciprocity failure mechanisms. The latent image centers formed from direct β-particle interactions are very stable and do not exhibit fading from reciprocity failure (Hamilton, 1977).

Detection of light emitted from scintillants is significantly improved by reducing the exposure temperature to –70° to –100°C. At these low temperatures, the latent image center forming in the silver grain becomes more stable, thus reducing latent-image fading. The detection sensitivity increases because the latent image accumulates with additional light exposure (see Table 10.11.2).

Intensifying-Screen Exposure

For use with an intensifying screen, the radioisotope must have sufficient energy to pass through autoradiography film and into the screen, which is located on the other side of film. ^{32}P and ^{125}I are the most common high-energy β and γ emitters that are exposed with intensifying screens.

In this procedure, under safelight conditions, the intensifying screen is placed in the cassette and a sheet of double-coated autoradiography film is placed on top of it. Next, a radiolabeled specimen is placed on top of the film with its radioactive side against the film and exposed at optimum exposure temperature—i.e., –70° to –100°C. At the end of the exposure, the cassette is warmed to ambient temperature before opening it. The film is removed carefully from the specimen after noting the orientation of the film sheet with respect to the specimen by bending a corner of the film towards the specimen or by marking a corner of the film with a no. 2 pencil. Then the film is photographically processed.

Inorganic phosphors in the intensifying screen convert ionizing radiation to blue light or blue and green light, depending on the type of intensifying screen. Generally, the spectral emission of the intensifying screen is matched to the spectral sensitivity of the film to produce optimal sensitivity. X-Omat AR film is spectrally matched with blue-light-emitting intensifying screens and BioMax MS film is spectrally matched to the blue- and green-light-emitting BioMax MS intensifying screen.

TranScreen Intensifying-Screen Exposure

In this method, ionizing radiation is amplified by placing a thin and uniform phosphor coating between the radiolabeled specimen and autoradiography film. The TranScreen ~LE intensifying screen is used to enhance the detection sensitivity with low-energy radioisotopes such as ^{3}H. The BioMax TranScreen LE intensifying screen folder is placed in the cassette and opened. Next, under safelight conditions, a sheet of BioMax MS film is placed inside the

Table 10.11.2 Sensitivity of X-Omat AR Film

Exposure procedure	Temperature (°C)	Sensitivity ($\mu Ci/cm^2$)[a]
Direct	20	8.00
Fluorographic	20	0.30
Fluorographic	–70	0.07

[a]Sensitivity is expressed as the amount of radioactivity required to produce a 0.10 optical density on X-Omat AR film (Eastman Kodak). [^{3}H]Glutaraldehyde bis bisulfite was spotted on Kodak Chromagram sheet and exposed to film for 24 hours (Laboratory and Research Products, 1993).

BCU IASI / CENTRAL UNIVERSITY LIBRARY

folder and the folder is closed on the film. The ^{3}H-radiolabeled specimen is placed on top of the folder and the light-tight cassette is closed and exposed at the optimum exposure temperature—between –70° and –80°C. At the end of the exposure, the cassette is warmed to ambient temperature before opening it. The film is carefully removed from the TranScreen folder after noting the orientation of the film sheet with respect to the specimen by bending a corner of the film towards the specimen or by marking a corner of the film with a no. 2 pencil. The film is then photographically processed.

BioMax MS film is spectrally matched to the BioMax TranScreen LE screen phosphor. The spectral output of the BioMax TranScreen LE is similar to BioMax MS intensifying screen.

Using the TranScreen LE to detect ^{3}H in a protein gel with western blot (10^{-4} cell fraction of a 10-atom label per protein of 30-kDa mass and a 10-μg gel load), the number of total discharges per mm^2 required to produce a 0.10 optical density above background was reduced ~16-fold, from 7.10×10^6 for direct exposure (Vizard et al., 1996) to 0.42×10^6 for TranScreen LE exposure.

PREFLASHING THE FILM

For very-low-light autoradiographic applications, such as with fluorography, detection may be improved by hypersensitizing the film and exposing it at reduced temperature. Hypersensitizing the film is achieved by preflashing it prior to autoradiographic light exposure. For this treatment to be beneficial, light must be the primary source of the autoradiographic film exposure, as in fluorography or intensifying-screen procedures. Preflashing does not provide any advantages with direct β-particle exposure procedures (see discussion of Direct Exposure).

Equipment needed for preflashing includes an electronic flash—e.g., Vivitar Auto 252 Electronic Flash or Sensitize Pre-Flash (Amersham), and light-filtering material. The Vivitar flash requires filtering the light through several sheets of Whatman 3-mm paper, a Kodak Wratten 22 gelatin filter, and a 3.0 neutral-density filter to control intensity. The electronic preflash unit from Amersham will most likely not require any additional filtering material.

The preflash density is the film-density difference between the exposed and the unexposed film. The optimum increase in preflash density is 0.10 to 0.15 optical density units. Preflashing produces the following various effects on photographic film.

1. It increases background density.
2. It lowers the contrast in the low-dose region of the film's characteristic curve (the characteristic film curve is a graph of film response to light exposure as measured by optical density).
3. It increases in net density above background.
4. It produces a more linear density-versus-exposure film response.

Information will be lost if preflashing is not in the optimal range. Thus, it is very important to develop a repeatable and effective preflashing procedure. Amersham provides detailed comprehensive instructions and calibration protocols with their preflashing unit (also see Laskey and Mills, 1975).

FILM-PROCESSING RECOMMENDATIONS

Autoradiography films are processed either in automated medical X-ray film processors or manually in 5-gallon tanks or small trays placed in a sink. With the exception of Direct Exposure Film (Eastman Kodak) and Hyperfilm-βmax (Amersham), most films are processable using the standard cycles on automated medical X-ray film processors. Direct Exposure Film and Hyperfilm-βmax must be processed manually. It is best to standardize film processing by using one processor and one method of processing, because film development in a manual method is slightly different than automated film development in a medical X-ray film processor.

Automated Processing

Automated medical X-ray film processing affords the most reliable and convenient processing of autoradiography films. Generally, the total processing time is <3 min. Besides the ease of use, most low-volume medical X-ray film processors produce high-quality autoradiograms.

Manual Processing

There are two different methods of manual processing for autoradiography films. The difference is primarily in the size and configuration of the containers used to hold the processing reagents. Both methods will yield high-quality autoradiograms. The deep-tank method involves several 5-gallon tanks surrounded by a temperature- and flow-controlled water bath. Film is attached to metal film hangers, and the

BIBLIOTECA CENTRAL UNIVERSITY LIBRARY

Table 10.11.3 Manual Processing Guidelines for Autoradiography Films[a]

Step	Solution	Agitation	Temperature (°C)	Time
Develop	Kodak GBX developer and replenisher	Do not agitate film sheet	20	5 min
Rinse	Kodak indicator stop bath or running water rinse	Use continuous agitation	16 to 24	30 sec
Fix	Kodak GBX fixer and replenisher	Intermittently agitate film 5 sec for every 60 sec	16 to 24	5 to 10 min
Wash	Clean running water with 2 changes every 15 min		16 to 24	10 min
Dry		Let stand in dust-free environment until dry	<50	

[a]For X-Omat AR, BioMax MR, BioMax MS, and BioMax Light films (Eastman Kodak; see *SUPPLIERS APPENDIX*).

hangers are moved from solution to solution by hand. The tray method involves at least three photographic processing trays on a bench or in a sink. The trays must be 2- to 3-cm longer and wider than the film sheet.

The processing chemistries, times, temperatures, and agitation specifications shown in Table 10.11.3 are a guide for manual processing of autoradiography film in either deep tanks or trays. The processing cycle indicated there is appropriate for most autoradiography, medical, and industrial films. Other types of developers and fixers may require different processing times and temperatures.

Autoradiographic imaging that is consistent and quantifiable relies upon consistent and reproducible processing. The developer and fixer solutions must be replenished daily or made fresh every 1 to 2 weeks to maintain solution activity level.

LITERATURE CITED

Hamilton, J.F. 1977. The mechanism of formation of the latent image. *In* The Theory of the Photographic Process, 4th ed. (T.H. James, ed.) pp. 105-145. Eastman Kodak Company, Rochester, N.Y.

Laboratory and Research Products. 1993. Autoradiographic Detection Principles. Eastman Kodak Company, Rochester, New York.

Laskey, R.A. and Mills, A.D. 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* 56:335-341.

Steinfeld, R., McLaughlin, W., Vizard, D., and Bundy, D. 1994. Advances in autoradiography: A new film brings quality improvements. *In* The Biotechnology Report pp. 147-149. Campden Publishing Ltd., London.

Vizard, D., Attwood, J., and McLaughlin, W. 1996. A universal intensifying screen system for enhanced detection of low- and high energy isotopes. *Am. Biotechnol. Lab.* 14(11):74-76.

INTERNET RESOURCES

http://www.kodak.com:80/daiHome/SIS/ksis.shtml

Site dedicated to scientific imaging systems and autoradiography products. Contains technical information, analysis software demos, and sales and service information for all Kodak Scientific Imaging Systems products.

Contributed by Daniel C. Bundy
Eastman Kodak Company
New Haven, Connecticut

BCU IAS CENTRAL UNIVERSITY LIBRARY

Digital Electrophoresis Analysis

Gel electrophoresis has become a ubiquitous method in molecular biology for separating biomolecules. This prominence is the result of several factors, including the robustness, speed, and potentially high throughput of the technique. The results of this method are traditionally documented using silver halide–based photography followed by manual interpretation. While this remains an excellent method for qualitative documentation of single-gel results, digital capture offers a number of significant advantages when documentation requires quantitation and sophisticated analysis. Digital images of gel electropherograms can be obtained rapidly using an image-capture device, and the images can be easily manipulated using image analysis software.

REASONS FOR DIGITAL DOCUMENTATION AND ANALYSIS

There are several reasons to consider digital documentation and analysis of electrophoresis results. These justifications can usually be categorized into issues of ease of handling, accuracy, reproducibility, and cost.

Ease of Handling

A major advantage of the digital revolution has been in storage and retrieval of information. Storage in notebooks and filing cabinets previously meant that searching for specific data or experiments was a tedious manual process. With digital information, modern search engines can quickly find specific information in a fraction of the time usually required for a manual search. Making backup copies of non-digital data can be difficult, expensive, and time-consuming since it requires copying, retyping, or photographic reproduction. Copies of digital data can be generated more easily and at reduced costs.

Manipulation of information is also easier when it is in a digital format. While the cut-and-paste analogy comes from physical documentation, it takes on a new perspective when applied digitally. Electrophoresis images can be resized, cropped, and inserted into reports. Data can be passed to spreadsheets and statistical packages for analysis and later insertion into notebooks and reports. These reports can be passed out via the Internet to colleagues throughout the world. A single individual can do all this in a few hours.

Digital analysis also provides an easier method for handling the data when comparing large numbers of results or large numbers of separate experiments. Research that requires comparing the banding patterns on 1000 gels containing 50 lanes each can be an undertaking of heroic proportions if the analysis is performed manually. Database software can dramatically speed the analysis and handle the more mundane tasks, leaving the researcher free to interpret the data.

Accuracy

The human eye is an extremely versatile measuring instrument. It can handle light intensities covering a range of nearly nine orders of magnitude and is sensitive to a fairly wide spectrum of light (Russ, 1995). Yet the eye cannot accurately and reproducibly quantitate density and patterns, nor can it deal with large numbers of bands or spots. Accuracy of measurement is a primary reason for using digital analysis on electrophoretically separated proteins and nucleic acids. Two categories of accuracy are key to digital analysis: positional accuracy, which is important for mobility determinations such as molecular weight, and quantitative accuracy.

Positional accuracy is based on both resolution of the recording medium and measuring accuracy. Silver halide–based recording has a theoretical resolution based on ~2000 imaging elements (silver grains) per inch. Measurement traditionally occurs using a ruler, with an accuracy of ~20 to 40 elements (50 to 100 elements per inch). In comparison, typical digital systems have 200 to 600 picture elements (pixels) per inch. The advantage that digital systems have is in measuring accuracy, which can occur at the level of a single imaging element.

Quantitative accuracy is also an issue. The amount of material represented by a band or spot is difficult to determine accurately from an image of a gel unless it is a digital image. On a digital image, the amount present is directly correlated with the derived volume of the band or spot—the volume is calculated using the intensity values of the pixels within the object.

Reproducibility

Any technique or measurement is only as good as its ability to be faithfully replicated. With software-defined routines, measurements

Contributed by Scott Medberry and Sean Gallagher
Current Protocols in Protein Science (1998) 10.12.1-10.12.14
Copyright © 1998 by John Wiley & Sons, Inc.

BYU LAW CENTRAL UNIVERSITY LIBRARY

are performed in the same manner every time. Allowing the computer to do repetitive tasks and complicated calculations minimizes the chance for individual errors. This does not imply that such measurements are correct, just that they are reproducible. An incorrect routine or algorithm can also invalidate data.

Cost

A consideration when evaluating any laboratory method is cost. Digital electrophoresis analysis equipment can be expensive. In many cases, however, it offers the only method for achieving acceptable analysis performance. In other cases, equal performance can be achieved using silver halide technology. However, traditional photography can also be expensive when the costs of consumable supplies such as film and developers, as well as other expensive requirements such as developing tanks and dark rooms, are included. Often, digital methods can be a good choice when all costs are considered.

KEY TERMS FOR IMAGING

There are several specialized terms encountered during digital image analysis. The most commonly encountered are contrast, brightness, gamma, saturation, resolution, and dynamic range. They describe controls on how the light detectors report a range of light intensities. Below is a brief description of each.

Contrast

Contrast describes the slope of the light intensity response curve. An increase in the contrast increases the slope of the curve. The result is a more detailed display over a narrowed range of intensities with less detail in the remaining portions of the intensity range. This is depicted in Figure 10.12.1A and 10.12.1B, where a normal, unadjusted image and a contrast-adjusted image are displayed, respectively. The contrast was increased on midrange intensity values in Figure 10.12.1B to highlight band intensity differences at the expense of background information. Images with a narrow range of informative intensities can benefit from increasing the contrast, since that effectively increases the scale and improves detection of minor differences in intensity. Contrast settings should be lowered if information is being lost outside of the contrast range. For example, in Figure 10.12.1B, loss of background information between peaks indicates that this image should not be used for quantitation.

Brightness

While brightness can have many different definitions, only one will be considered here. Brightness shifts the light intensity response curve without changing its slope as is shown in Figure 10.12.1C. Another name for brightness is black level, since it is commonly used to control the number of black picture elements (pixels) in an image. Incorrect brightness levels can lead either to high background and potential image saturation or, as is illustrated in Figure 10.12.1C, to a total loss of background information and partial loss of band information.

Gamma

Nonlinear corrections are often applied to images to compensate for how the eye perceives changes in intensity, how display devices reproduce images, or both. The most common correction is an exponential one, with the exponent in the equation termed the gamma. A typical gamma value for camera-based systems is 0.45 to 0.50, and is illustrated in Figure 10.12.1D. This is a compromise value that compensates for the 2.2 to 2.5 gamma present in most video monitors and the print dynamics of most printers. Since it is a nonlinear correction, special care must be taken if quantitation is desired. Unless directed to by the manufacturer, gamma values other than 1.0 should be avoided when quantitating. More information on gamma correction can be found on Poynton's Gamma FAQ (*www.inforamp.net/~poynton/Poynton-color.html*).

Dynamic Range

Dynamic range describes the breadth of intensity values detectable by a system and is usually expressed in logarithmic terms such as orders of magnitude, decades, or optical density (OD) units. A large dynamic range is important when trying to quantitate over a wide range of concentrations. The most accurate quantitation occurs in the linear part of the dynamic range, which is usually not the complete dynamic range of the system. An additional consideration is the dynamic range of the visualization method. Many popular visualization methods have linear dynamic ranges of 1 to 2.5 orders of magnitude. An imaging system with greater dynamic range analyzing the results of such a visualization method will not improve the dynamic range.

BCU IASH GENERAL UNIVERSITY LIBRARY

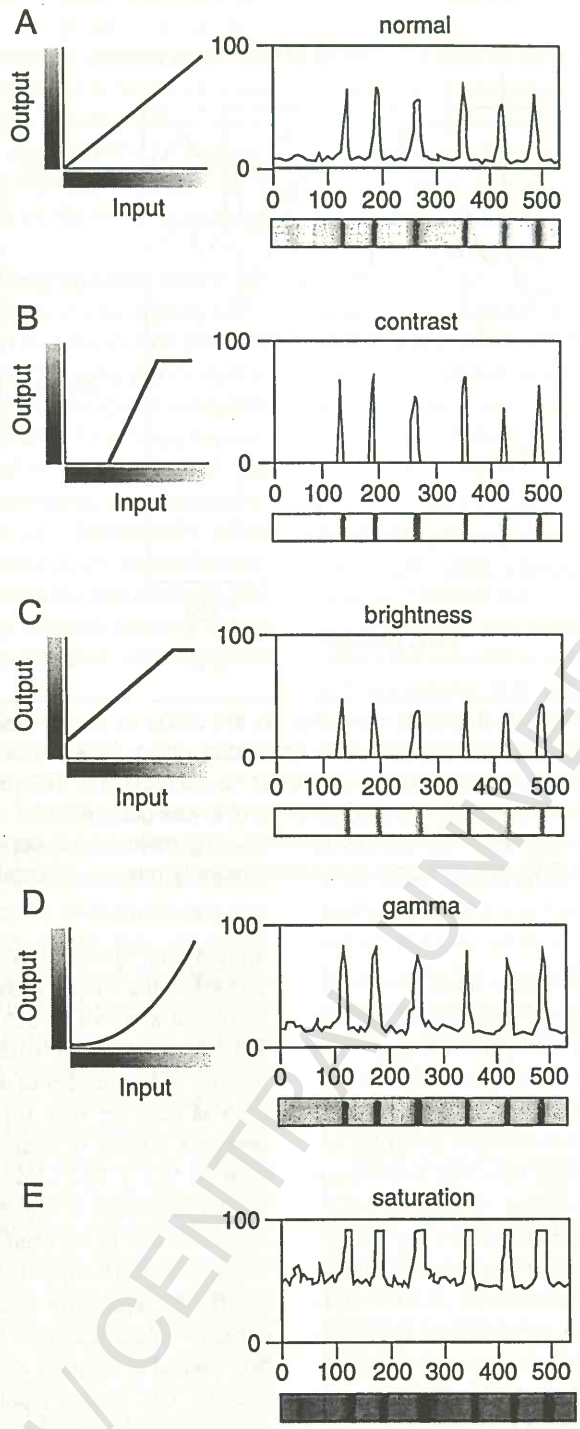


Figure 10.12.1 Examples of how altering image capture settings affects the image and the analysis. The graph on the left displays the light intensity response curve used for image capture while the image and resulting lane profile on the right display how the setting affects the image. The lane profile displays pixel position versus normalized pixel intensity (**A**). In this case, the output has not been altered, giving a straight line with a slope of 1 on the response curve. (**B**) The image acquisition was adjusted to increase the contrast of the displayed image. Although useful for images with a narrow range of informative intensity values, increasing the contrast can lead to a loss of low and high values. (**C**) Decreasing the brightness reduces peak values but also leads to a loss of the weak bands and original background. (**D**) Gamma adjusts raw data to appear more visually accurate. Note that this leads to a loss of fidelity between the adjusted image and the original. (**E**) Saturation indicates that the detector is reporting its maximum value or that the dynamic range for the visualization method has been exceeded.

ASU / CENTRAL UNIVERSITY LIBRARY

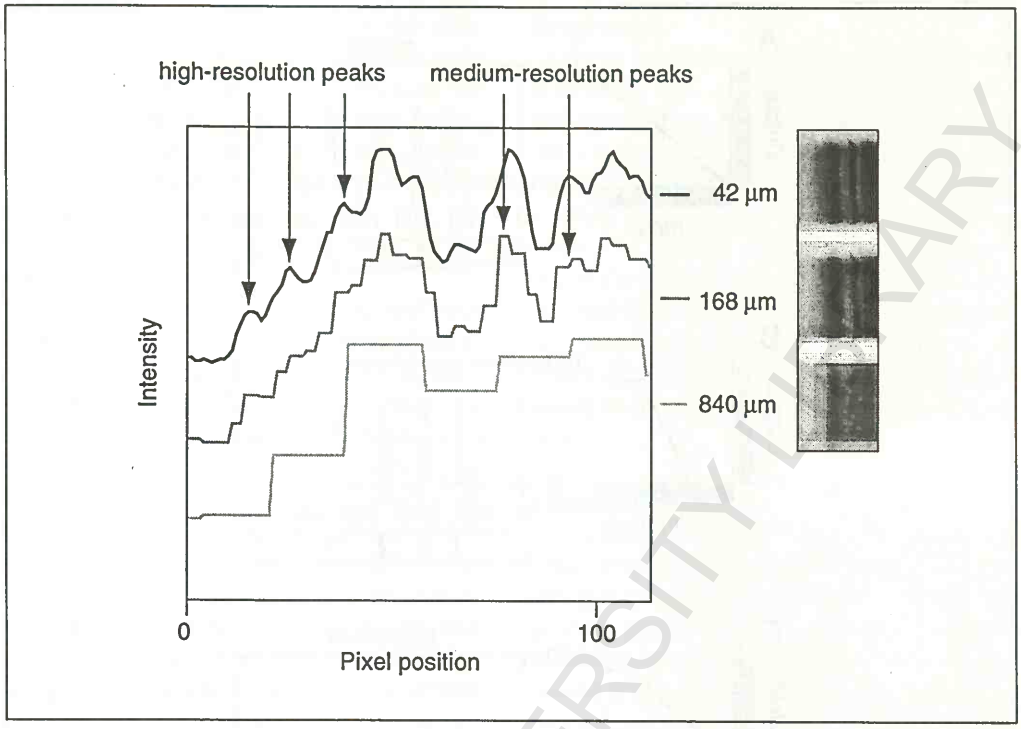


Figure 10.12.2 The effect of spatial resolution on the ability to detect closely spaced objects. Whole-cell protein lysates from *E. coli* were separated using SDS-PAGE and visualized with Coomassie blue staining. An image was captured at 42 µm (600 dpi), 168 µm (150 dpi), and 840 µm (30 dpi) from a segment of the lane, and a lane profile was generated for each image. The lane profiles have offset intensities to allow for comparison. Only major bands can be detected with the low-resolution image (840 µm); at higher resolutions more bands are detectable.

Saturation

Saturation occurs when a detector or visualization method receives input levels beyond the maximum end of the dynamic range. This results in a loss of detail and quantitative information from those data points that are saturated. For fluorescent and luminescent samples, reduction in the sampling time can sometimes correct saturation problems. Optical density-based visualization techniques can also generate saturated images, as is illustrated in Figure 10.12.1E; this can sometimes be avoided with longer sampling times or increased detection source intensities. More often, it will be necessary to perform another electrophoresis with more dilute samples or to alter the visualization process to generate a less optically dense material.

Resolution

Resolution is the ability of a system to distinguish between two closely placed or similar objects. Three types of resolution are important for analysis—spatial resolution, intensity resolution, and technique-dependent resolution.

Spatial resolution is the ability to detect two closely placed objects in one-, two-, or three-dimensional space. It is most accurately described as the closest distance two objects can be placed and still be detected as separate objects. In practice, it is often defined nominally in terms of the number of detectors per unit area such as dots per inch (dpi) or the number of detectors present in total or in each dimension such as 512 × 512 (262,144 total detectors). Actual resolution is less than half the nominal resolution due to the need for two detectors for every resolvable object (one for the object and one for the separation space) and the effects of optical resolution. Figure 10.12.2 demonstrates how spatial resolution can affect detection of objects. The 42-µm resolution image allows detection of closely spaced bands, the 168-µm resolution image detects fewer bands, and the 840-µm resolution image detects only major bands. For instruments with on-line detection systems, a pseudo spatial resolution is often reported in units of time from the start of the separation or the time interval between two objects crossing the detection path.

Intensity resolution is the ability to identify small changes in intensity. It is a function of both the dynamic range of a detector and the number of potential values that detector can

BCU IAST CENTRAL UNIVERSITY LIBRARY

report. Greater dynamic range decreases the intensity resolution of a given detector. The number of potential values a detector reports is described by its bit depth. An 8-bit detector can report 256 (2^8) different possible values, while a 12-bit detector can report 4,096 (2^{12}) values, and a 16-bit detector can report 65,536 (2^{16}) values. The higher the bit depth, the greater the intensity resolution.

Technique-dependent resolution directly affects the spatial and intensity resolution. Electrophoretic separation techniques that generate overlapping objects or that have object separation distances shorter than the spatial resolution will fail to provide reliable data. Many factors, including the amount of sample loaded, gel pore size, buffer constituents, and electrophoresis field strengths, can dramatically affect separation and resolution of biomolecules. Likewise, detection methods that can only generate a small range of discrete intensity values will not benefit from systems with improved intensity resolution.

IMAGE CAPTURE

Devices

Capturing digital images involves a detection beam or source, a sensor for that beam or source, and some method of assembling a two-dimensional image from the data generated. Most systems use a light source for detection. The light wavelengths used range from ultraviolet (UV) to infrared (IR) and can be broad spectrum or narrow wavelength. Broad-spectrum detection is more versatile since it can often be used for more than one detection wavelength. However, when compared to narrow-wavelength sources such as lasers, broad-spectrum detection suffers from reduced sensitivity and reduced dynamic range. While many types of light sensors have been used, including charge-coupled devices (CCDs), charge-injection devices (CIDs), and photon multiplier tubes (PMTs), technology advances in CCDs have led to their dominance. CCDs are semiconductor imaging devices that convert photons into charge. This charge is then read and converted into a digital format via an analog-to-digital converter (ADC).

The method of image assembly depends on the light source and detector geometry. One method is to capture the image all at once using a two-dimensionally arrayed CCD detector similar to the detectors found in digital and video cameras. Typically a camera-type sensor is paired with a light source that evenly illuminates the sample. This same sensor is often used with fluorescent and chemiluminescent detection methods, as its ability to detect light continuously over the entire sample reduces image capture times. Another method of image assembly is to capture the image a line at a time. This typically involves a linearly arrayed CCD scanning slowly across the sample in conjunction with the detection beam of light. The data from each line is then compiled into a composite image. Spatial resolution in this method can be significantly better on large-format samples than the resolution of a camera-based system. This method is also advantageous when OD-based detection is used, since the more focused light beam is usually of higher intensity and can penetrate denser material. A third method of image assembly is to use a point light source and single-element detector on each point on a sample. The image is then compiled from each point sampled. This method is slower than the others but can offer extremely high resolution and sensitivity. A fourth commonly encountered method is that of generating a pseudoimage of electrophoresis results through the use of a finish line type of detection system. This is comprised of a light source positioned at the bottom of the gel (i.e., the end opposite of the site of sample loading) and light detectors positioned next to each lane to detect the transmitted light or emitted fluorescence. A lane trace is generated using time on the *x* axis and light intensity on the *y* axis. The pseudoimage is then generated from this data (Sutherland et al., 1987).

Capture Process

Prior to image capture, electrophoretic separation and any visualization steps are performed. To calibrate the separation process, standards are usually run at the edges of the gel and often at internal positions. If quantitation of specific proteins or nucleic acids is to be performed, a dilution series of standards with similar properties to the experimental samples should also be included. After separation, the protein or nucleic acid is visualized if necessary. Visualization can include binding of a fluorochrome or chromophore such as Coomassie blue, precipitation of metal ions such as copper, silver, or gold, enzymatic reactions, and exposure of film or phosphor screens to radiant sources. These methods can be grouped based on the type of detection into optical density, fluorescence, chemiluminescence, and radioactivity. The suitability of popular detection devices with these methods

BOURNEMOUTH UNIVERSITY LIBRARY

Table 10.12.1 Compatability of Popular Image-Capture Devices with Common Visualization Methods[a]

Visualization method	Image-capture device				
	Silver halide photography	CCD camera	Desktop scanner	Storage phosphor	Fluorescent scanner
Optical density[b]	+	+	++	−	−
Fluorescence	+	+	−	−	++
Chemiluminescence	+	++	−	±	−
Radioactivity	+	−	−	++	−

[a]The device with the highest sensitivity and greatest dynamic range for a visualization method is marked with a ++, other devices that can detect this visualization method are indicated with a +, and devices that are not suitable for a visualization method are indicated with a −. A ± indicates that only some devices of this type can be used with this visualization method.

[b]Optical density methods include Coomassie blue staining.

is described in Table 10.12.1. Once visualization has occurred, image capture consists of the following steps: previewing the image while adjusting capture parameters, capturing the image, and saving the image for later analysis.

During the preview process, capture parameters are optimized for data content and for ease and rapidity of later processing steps. Typically, the first step is to place the sample so that when the image is captured, the rectangular edges of the gel are horizontal and vertical on the monitor and any lanes are either horizontal or vertical. Since band and spot detection will be much easier if the image is properly oriented, this eliminates the need to later rotate the image digitally. Image rotation is time-consuming and can result in spatial linearity errors (a change in the size and shape of objects in image) caused by rectangular image-capture device geometries. The next step for camera-based systems is to adjust magnification and to focus the sample image. For thicker samples, it might be necessary to reduce the aperture on camera-based systems to get a sufficient depth of field to focus the entire sample. Often at this point image imperfections—e.g., dust, liquid, or other foreign objects that will detract from later analysis—are detected, and they need to be removed. Next, image intensity is set. Within the area of interest on the image, band or spot peaks should have values less than the maximum saturated value, and the background should have nonzero values. This is usually accomplished through adjustment of the light-source intensity or the sensor signal integration. If the device allows precapture optimization of other parameters such as spatial resolution, contrast, brightness, gain, or gamma, they are adjusted next. Note that this only applies to controls that affect the response of the sensor or processing of the image prior to a data reduction step and not to controls that affect the image at later stages. The latter process can enhance visualization of specific features but is best left to adjustments in look-up tables (LUTs) in later analysis steps rather than during image capture, since there is a risk of data loss during postacquisition image processing. LUTs are indexed palettes or tables where each index value corresponds to color or gray-scale intensity values present in an image. Many image analysis programs alter LUTs instead of image values directly, since it both is faster and does not change the original image data.

Once all the capture parameters are optimized, the image capture process is initiated. This might take less than a second for images captured with camera-based detectors and up to hours for scanning single-point detectors. When the image has been captured, it should be carefully examined for content. It should fully capture the area of interest and the parameters should have been set so that all necessary information is detectable. Furthermore, it should be in a form that will allow for easy analysis. Extra time spent optimizing the capture parameters will often result in a reduction in total image analysis time and in an increase in data quality. When the best possible image has been captured, it often contains information outside the area of interest. While this is unlikely to cause problems with later analysis, it is often advantageous to crop the image so that the only portion that is saved contains the area of interest. This reduces the amount of disk space necessary to store the image, and the image usually will load and analyze faster with the analysis software.

RCU IAS CENTRAL UNIVERSITY LIBRARY

The last step in image capture is to save the image. Several options are available at this point, including choosing which location to save the image at, what file type or format to use, and whether to use some form of compression.

The location where the image is saved is not as trivial a question as it might seem if the image will need to be transferred to another computer at some point. File sizes can easily exceed 15 megabytes on high-resolution images. This is a manageable size for hard drives but exceeds current floppy drive sizes by an order of magnitude. There are software utilities available that will subdivide files into disk-size chunks and then reassemble them at the next computer, but this is an inconvenient and slow method. If the computer used to help capture the image is connected to a network, the image files can easily be transferred this way or potentially saved on a central server. Alternatively, several types of high-capacity removable media are available (e.g., Zip or Jazz). This usually requires the installation of additional hardware onto two or more computers but does make backing up data easier.

Since image files can be very large, compression techniques are sometimes used to reduce disk space requirements. Compression algorithms use several methods, typically by replacing frequent or repetitive values or patterns with smaller reference values and by replacing pixel values with the smaller difference values describing the change in adjacent pixels. When the file is later decompressed, the compressed values are then replaced with the original information. Not all images compress equally, with simple images containing mostly repetitive motifs compressible by ≥90%, while complex images will benefit much less from compression. Because compression is a much slower method of saving files and will not benefit every file, compression is not used to save all files. Several different forms of compression are available but are separable into two main classes, lossless and lossy. Lossless methods faithfully and completely restore the image when it is decompressed (no loss of data) but offer only moderate file compression; compression values range from ~10% to 90%, depending on the image. Examples of lossless compression include Huffman coding (Huffman, 1952), RLE (Run Length Encoding), and LZW (Lempel, Ziv, and Welch; Welch, 1984). In comparison, lossy methods such as JPEG (Joint Photographic Experts Group), MPEG (Moving Picture Experts Group), or fractal compression schemes can reach compression values of ≥98% (Russ, 1995). The trade-off is that not all information from the original file is recovered during decompression. Lossy compression is sometimes necessary for applications with extremely large image files such as real time video capture, but it usually represents an unacceptable loss of data if used with electrophoresis image capture.

Many different file types have been developed to store digital images. Some of these file types are proprietary or hardware specific. For example, PICT is a Macintosh format and BMP is a PC-compatible format. Each file type has its own structure. Some types do not allow compression, for others it is optional, and for some it is mandatory. File types vary in the types of images they support, particularly in the number of colors or gray levels. Below is a brief description of a few of the more prevalent file types.

TIFF (Tagged-Image File Format) is one of the most commonly used formats. It is particularly versatile since it is an open format that can be modified for specific applications. One reason for its versatility is the ability to attach or tag data to the image. The tags can include information such as optical density calibration, resolution, experimenter, date of capture, and any other data that the application software supports. TIFF images can be monochrome, 4-, 8-, or 16-bit gray-scale, or one of many color-image formats. Compression is optional, with LZW, RLE, and JPEG often supported (Russ, 1995). Since TIFF is supported by both Macintosh and PC computers, it is a good choice for multiple-platform environments. The versatility of TIFF can also be a weakness. Since there are many different tagging schemes and since not all programs support all possible compression and color schemes, it is sometimes not possible for one program to access the information in a TIFF file generated by a different program.

GIF (Graphics Interchange Format) is a file format that is widely encountered on the Internet due to its compactness and standardization. Its compactness is attributable to a mandatory modified LZW compression. Another feature of GIF is the use of a LUT to index the values in the image. One interesting ability of GIF is that it supports storing multiple images within a single file. This can offer some advantages for applications such as time-lapse image capture. A GIF image can contain no

BRANDEIS UNIVERSITY LIBRARY

more than 256 individual color or gray levels and does not support intensity resolutions higher than 8 bit. In addition, since the image is implemented as a LUT, it also is not a true gray-scale image. Due to these limitations and others, alternative formats such as PNG have been developed to replace GIF.

PICT is a file format and graphics metafile language (it contains commands that can be played back to recreate an image) designed for the Apple Macintosh. It can contain both bitmap images and vector-based objects such as polygons and fonts. It supports a ≤256-gray-level LUT, and monochrome images can be RLE compressed. Because it only offers a 256-gray-level LUT, it has the same weaknesses that GIF does with true gray-scale and high-intensity-resolution images. In addition, any vector objects in the image are difficult to translate on a PC since they are designed to be interpreted by Macintosh QuickDraw routines.

BMP is the native bitmap file format present on Windows-based PCs. It supports 2-, 16-, 256-, or 16-million-level images. With images of ≤256 gray levels, it implements a LUT, while the highest-resolution image is implemented directly. RLE compression is optional for 16- and 256-gray-level images. Since compression is prohibited on 16 million-gray-level images and there is no intermediate level supported beyond 256 levels, BMP is not a good choice for images with high-intensity resolution requirements.

ANALYSIS

Once the image has been captured, the data needs to be analyzed and distilled into information about the results of the electrophoresis experiment. Through the use of standards and experimenter input, this software-driven process can estimate mass and quantity of objects in an image and detect relationships between objects within one image and between similar images. The type of software used depends on the analysis to be performed. Images from single electrophonetic separations are examined by one-dimensional analysis software optimized for lane-based band detection. Images from two-dimensional electrophoresis are best handled by specific programs designed to detect spots and to assign two mobility values and a quantity value to the spot. After the initial characterization of bands and spots, comparisons are often made between bands or spots from different experiments through the use of database programs and matching algorithms.

Software for One-Dimensional Analysis

Lane positioning

For one-dimensional analysis, the first activity is to detect the lanes on the image. One of three different methods is commonly employed for this. For images with straight, well-defined lanes with a large number of bands, automatic lane-detection algorithms can quickly and accurately place the lanes. On images with very well-defined lanes, such as pseudoimages from finish-line type electrophoresis equipment, automated lane calling based on image position is possible. For images with "smiling," bent, or irregular lanes, manual positioning of the lanes is often the fastest and most accurate method of lane definition. Regardless of the method of identifying the lanes, the lane boundaries need to be carefully set for accurate quantitation and mass determinations. Lane widths should be wide enough so that the entire area of all bands in that lane are included, but they should not be so wide as to include bands from adjacent lanes. To accomplish this, curved or bent lanes might need to be used in order to follow the electrophoresis lane pattern. Lane length and position also must be adjusted as necessary so that all bands of interest are included. If mass determinations are necessary, the sample loading point should probably also be included in the lane or be the start of the lane. At this point, lines of equal mobility (often called R_f or iso-molecular-weight lines) are added to the image as necessary. These lines allow for correction of lane-to-lane deviations in the mobility of reference bands and generate more accurate measurements of mass. A similar form of correction is also possible for within-lane correction of mobilities. This correction is important for accurate detection and quantitation of closely spaced bands.

Band detection

Once the lanes have been defined, the bands present in each lane need to be detected. There are many methods for detecting bands. One method is to systematically scan the lane profile from one end to the other, identifying regions of local maxima as bands. Another common method is to use first- and second-order derivatives of the lane image or lane profile in order to find inflection points in the change of slope in pixel intensity values (Patton, 1995). Regardless of the method used, it is often necessary to alter the search parameters so that they perform reliably under a given experimental condition.

BCU IASE CENTRAL UNIVERSITY LIBRARY

Typical search parameters include ones for detection sensitivity, smoothing, minimum interband gaps, and minimum or maximum band peak size. Smoothing reduces the number of bands detected due to noise in the image. A minimum interband gap is often used to avoid detection of false secondary bands on the shoulders of primary bands. Limitations on peak sizes, especially for within-lane comparison to the largest band's peak, can be a useful way to allow sensitive detection of bands in underloaded lanes without detecting false bands in overloaded lanes.

Band edges are often detected in addition to band peaks in order to further define bands or to quantitate band amounts. This can be accomplished by using local minima, derivatives of the lane profile, or fixed parameters such as image distances or a percentage of band peak height. The band edges can be applied as edges perpendicular to the long axis of the lane or as a contour of equal intensity circling the band. The perpendicular method is advantageous for bands with uneven distribution of material across the face of the band, while the latter method is better for "smiling" or misshapen bands.

Background subtraction

With nearly all electrophoresis procedures, the most informative images have a low level of signal intensity at each pixel that does not result from protein or nucleic acid. Instead, this background intensity is attributable to the gel medium, the visualization method, electronic noise, and other factors. Since this background tends to be nonuniformly distributed throughout the image, failure to subtract it can make band detection and quantitation less accurate. Many methods of background subtraction are possible. Sometimes it is possible to generate a second image under conditions that do not detect the protein or nucleic acid. The second image is then digitally subtracted from the data-containing image to remove the background. More often, background information must be obtained from a single image. If the background varies uniformly across the image, a line that crosses the variation can be defined at a point where no bands are present. The intensity values at each point on the line can be used as the background value for the pixels perpendicular to the line at that point. Commonly, background is also present as variations in intensity along the long axis within each lane. One simple method is to take the lowest point in the lane profile as the background. Another method is to use an average value of the edge of each band as the background for that band. More complicated methods such as valley-to-valley and rolling-disk use local minima points in the lane profile to define a variable background along the length of the lane. Because there can be many different causes and distributions of background, no single method of background determination can be recommended for all experiments.

Characterization

Once lanes and bands have been detected, it is possible to interpret the mobility of the nucleic acid or protein bands. Depending on the method of electrophoretic separation, information on mass (length or size), pI, or relative mobility (R_f) can be inferred from mobility information. The mobility is characterized using a standard curve with internal standards of known properties. The type of curve depends on several parameters. By definition, with R_f-based separation, a linear first-order curve is used since it represents the linear relationship between mobility and R_f. Similarly, pI and mobility are generally linear in isoelectric focusing separations. For separations based on size, a curve generated from mobility versus the log of the molecular weight provides a relatively good fit as measured by the correlation coefficient (R^2). Several other curves have been suggested for size-based separations, including modified hyperbolic curves and curves of mobility versus (molecular weight)$^{2/3}$ that have good correlation coefficients (Plikaytis et al., 1986). In some cases, no single curve equation can adequately represent the data, and methods of fitting smooth contiguous curves using only neighboring points, such as a Lagrange or spline fit (described in Hamming, 1973), are necessary. This is most common for size separations with a very large range of separation sizes and with nonlinear gradient gels. Care must be taken with multiple-curve techniques since they rely on only a few data points for any one part of the composite curve, and outlying data points can drastically affect the outcome.

For size and R_f determination, a uniform position must be found in each lane as a point from which to measure the mobility of each band. Many software analysis packages use the end of the lane as the measuring start point, so for them it is important to position each lane start point at an iso-molecular-weight or iso-R_f point. A convenient point is the well or sample-loading position since it is usually easily de-

B.C.U. „M. EMINESCU" IAŞI
CHI . IE

BIBLIOTECA CENTRALĂ UNIVERSITARĂ

tectable and at an equal mobility position in each lane. A consistent point on each band must also be chosen to measure mobility. A band's peak is easily defined in digital image analysis and is commonly used. Since peak positions are harder to detect visually than edges on silver halide images, the leading band edge is sometimes used when comparing digital results with silver halide–based results.

Once lanes and bands have been detected, it is also possible to quantify the amount or at least relative amount of nucleic acid or protein present in each band. The amount in a band is related to the sum total of the intensity values of each pixel subtracted by the background value for each pixel in a band. For absorptively detected bands, intensity values are converted to OD values. The total value that is calculated is equivalent to the volume of the band and can be directly compared to other bands that are within the linear range for the visualization method. If standards of known amounts are loaded onto the same gel, they can be used to generate a standard curve that converts band volume into standard units such as micrograms. For greatest accuracy, it is important to be able to generate multiple standard curves when using visualization methods, such as Coomassie blue staining, that are affected by band or spot composition.

Quantitation becomes more complicated when bands are not fully resolved. In this case, material from one band is contributing to the volume measurement of an adjacent band and vice versa. The simplest method for handling this is to partition into each band only the volume within its edges. Alternatively, a Gaussian curve can be fitted to each band and the volume contained within the curve used to estimate the amount of the band. Since most electrophoresis bands have a pronounced skew towards the leading edge of the band, modified Gaussian curves have also been used (Smith and Thomas, 1990). In either case, the curve-fitting process is calculation intensive and can significantly increase analysis times for images with many bands.

Software for Two-Dimensional Analysis

In two-dimensional analysis, the first-dimension separation is performed in a single column or lane and then a second separation is performed perpendicular to the first. The result after visualization is a rectangular image of up to 10,000 spots. The most common two-dimensional gel type is one in which protein is separated first by apparent pI and second by molecular weight, although two-dimensional separation of nucleic acids is also possible. While many of the concepts and analysis techniques used with one-dimensional gels are applicable to two-dimensional gels, the complex nature of most two-dimensional gels requires somewhat different methodology. For example, spots are more difficult to detect since they are not conveniently arranged in lanes and can vary more in shape and overlap than bands. In addition, two-dimensional experiments usually require some method of comparing between two images, whereas one-dimensional images usually contain all of the information from an experiment.

Spot detection

Probably the most difficult aspect of two-dimensional analysis is efficient and accurate spot detection. If it is incorrectly done, it can lead to hours of manual editing. Due to the complexity and computational intensity of some algorithms, the detection process itself can last hours on relatively fast desktop computers. One theoretically effective but computationally intense method is to treat the image as essentially a three-dimensional image with spots treated as hills and background as valleys. A large number of Gaussian curves are then combined to describe the topology of the image. Many other methods make use of a digital-imaging technique known as filtering. In essence, filtering is a way to weight the value of a pixel and its neighbors in order to generate a new value for a pixel. By passing a filter across an image pixel by pixel, a secondary image is generated. Filters can be designed for many tasks, including sharpening an image or removing high-frequency noise. Filters can also be generated to help detect spots by making images that are first and second derivatives of the original image. The derivative images indicate inflection points in the intensity pattern and can be used to detect spot centers and edges. In a different method, called thresholding, filters can be used to detect the edges of objects. Instead of looking for inflection points, threshold filters identify intensities above a set level or at ratios between central and edge pixels above a set value. Since the edges on two-dimensional spots tend to be diffuse, sharpening filters are sometimes used prior to the thresholding filter. In some cases, multiple techniques are used to detect spots (Glasbey and Horgan, 1994).

BCU IAS CENTRAL UNIVERSITY LIBRARY

Unlike one-dimensional detection, detection on images of two-dimensional experiments usually requires secondary processing to get acceptable performance. One example of a secondary process is to discard spots with sizes below a set minimum or above a set maximum. Another is to analyze spots that are oval for possible splitting into two spots. Even after secondary processing, it is likely that a small amount of manual editing will be necessary. When manually editing an image, care must be taken to use as objective criteria as possible, especially if two or more images are to be matched and spot volumes compared.

Characterization

In a two-dimensional system, determination of protein or nucleic acid mobility is complicated by the fact that there are two mobilities to account for and that the second-dimension separation tends to make estimation of the separation which occurred in the first dimension more difficult. One method for dealing with this is to have a series of markers in the sample that, after both separations are completed, are evenly distributed within the gel and image. It is also possible to estimate separation characteristics from calibration points located at the periphery of the gel. For example, distance measurements can be used to pass calibration data from the first dimension separation, and standards can be separated at the ends of the gel to calibrate for mobility in the second dimension. Regardless of the method used, in many instances, a series of related images will be examined and similar spots in each image will be matched. When this occurs, it is possible to calibrate one image and then pass the calibration information via the matches to the related images.

Quantity determination is similar in many regards to that which occurs in one-dimensional analysis, but there are some differences. If spot edges are detected, a simple method of determining spot volume is to take the sum of the intensity value of each pixel in the spot reduced by a background intensity value. Multiple Gaussian curves can also be fitted to the spot to approximate the volume (Garrels, 1989). More difficult is attempting to compensate for a skewed distribution in a size-separating dimension while trying to use a regular Gaussian fit for a pI separation, such as is encountered with the most common form of two-dimensional protein separations. The distribution of background makes quantitation more difficult in two-dimensional gels. There is no lane-dependent component, so it is necessary to use other methods such as image stripes, finding local minimum values or using values derived from the spot edges to determine background values.

Matching

Matching is the process in which proteins or nucleic acids with similar separation properties are linked or clustered together. Matching can occur within one image or between multiple images as long as a frame of reference is established. Matching allows for comparisons between samples. It also makes annotation and data entry easier, since if one spot or band is matched to others and is characterized or annotated, this information is easily passed to all the other matches. An underlying assumption of matching is that objects with similar separation properties are actually similar. Care must be taken to confirm the identity of matched spots or bands by other methods on critical experiments.

A simple form of matching is to link bands or spots at similar positions on the gel images. This works well when separation and imaging conditions are uniform. This is very seldom the case, since slight differences in the electrophoresis, visualization, and imaging conditions across a gel and between gels generates incorrect matching with this method. Since bands on one-dimensional gels are relatively easy to calibrate for mobility, matching can occur along contours of equal mobility. This dramatically decreases but does not eliminate the variability in detecting similar bands. Much of the remaining variability can be attributed to calibration errors. This error can often be compensated for by allowing a small tolerance in mobility values in determining whether a band is matched or not. Because of the difficulties in calibrating mobility in two-dimensional gels, it is often more practical to use matched spots for calibrating mobility than vice versa. Spot matching between two two-dimensional images starts with finding a small number of landmark spots that are used as seeds for subsequent matches (Appel et al., 1991; Monardo et al., 1994). There are many methods for finding the landmarks in both images, including finding the highest-intensity spots, finding spots in unique clusters, and manual positioning. The most common procedure from this point is to derive a vector that describes the direction and extent of the path from one matched spot to the other when the two images are superimposed. The vector is used as the basis for finding more matches near the landmark matches. To allow

BOSTON UNIVERSITY CENTRAL LIBRARY

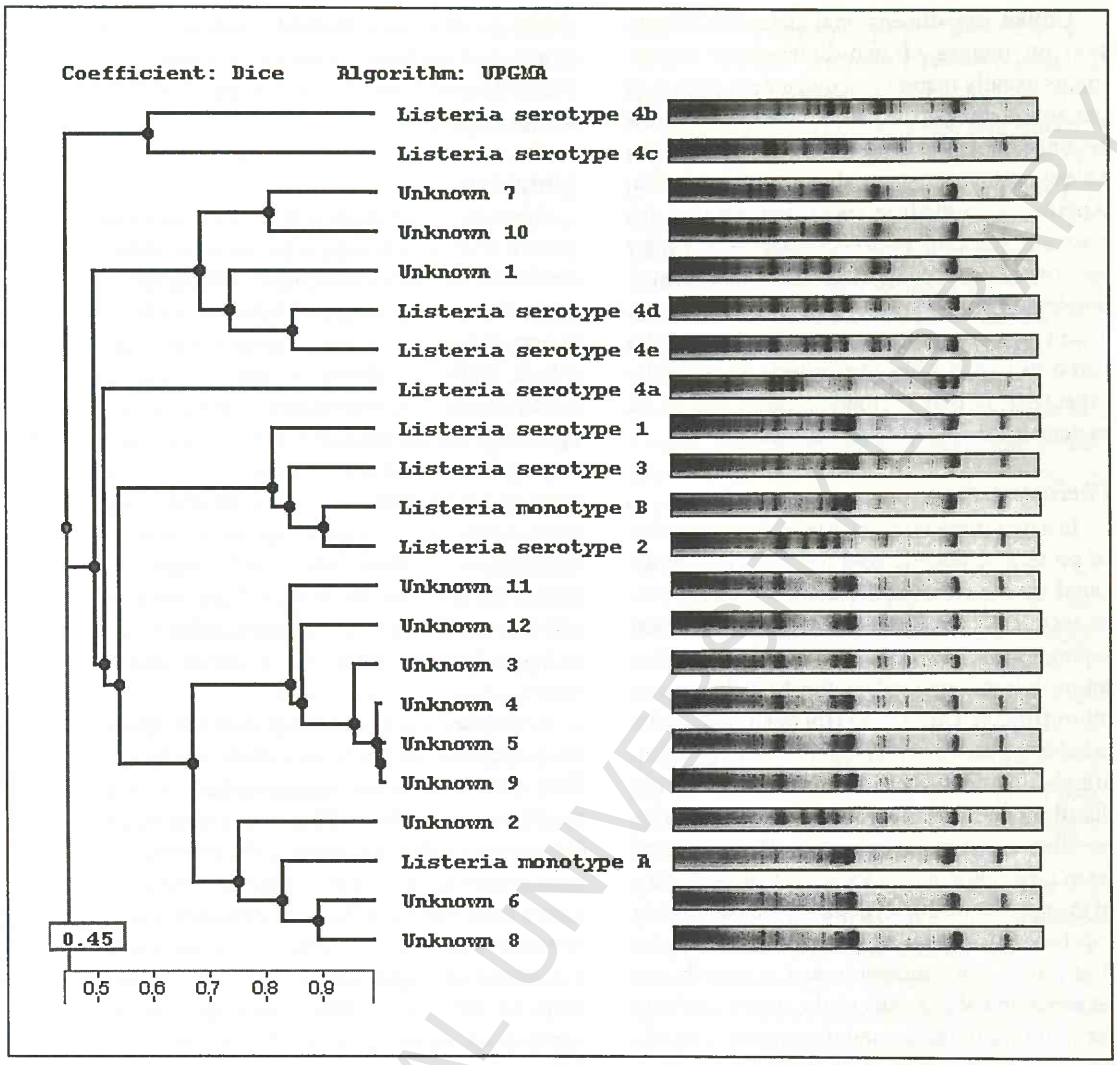


Figure 10.12.3 Example of a dendrogram generated from similarity data on band matching between lanes. DNA samples from 22 isolates of *Listeria* were subjected to Random Amplification of Polymorphic DNA (RAPD) analysis and the resulting electrophoresis image was analyzed with ImageMaster software (Amersham Pharmacia Biotech). Clustering was performed using the Dice coefficient with a tree structure based on the Unweighted Pair-Group Method using Arithmetic Averages (UPGMA). Similarity values between isolates can be determined by locating the node that connects the isolates and reading the value from the scale on the lower left edge of the dendrogram.

for error, the area within a small radius is searched extending from the end of the vector. Once another match is found, its vector is computed and used as the starting point for finding neighboring matches. From this progression, the entire gel is matched. If all vectors are displayed graphically when matching is complete, questionable matches can often be identified as vectors that are significantly different from neighboring vectors.

One specialized use of matching is as an estimator of the similarity and potential genetic relatedness of organisms. For example, on a one-dimensional gel image, a ratio of the matched to unmatched bands for each pairwise combination of lanes can be calculated. This ratio can be used as an indicator of similarity, with values near 1 indicating a pair of highly similar lanes, and values near zero indicating very dissimilar lanes. Assuming that the contents of the lanes are valid samples of the originating organism's genetic makeup, the information on lane similarity can be converted to estimates of genetic similarity. A convenient way to display this similarity data graphically is to generate a dendrogram with similar objects close to each other and less similar objects more distantly placed. An example of such a dendrogram is presented in Figure 10.12.3, where samples from *Listeria* isolates are arranged based on banding pattern.

BCU IACT CENTRAL UNIVERSITY LIBRARY

Databases

In many cases, image analysis is not the last step in the process. The image and analysis data need to be archived in a searchable format. There may be a need to analyze the data from multiple experiments conducted at different sites or in laboratories around the world. Bioinformatic links to diverse data sources might be desired to help develop a unified understanding of the biology behind particular phenomena. When these situations arise, database programs can be utilized to store, link, and search image analysis results.

As the number of images that are captured and analyzed grows, it becomes increasingly more difficult to find particular information from the large number of files that are stored. Relatively simple databases can be used if the major requirement is to find previously analyzed images and associated data. Such databases often display a miniaturized version of each image to aid in visual scanning for the file as well as simple searching for image-specific information such as date of analysis, file name, or other information that was entered at the time of image capture. More powerful database products are also available that can perform complex searches on data generated during the analysis. For an example, a search on a two-dimensional database might include finding proteins exhibiting a specific expression profile and having a molecular weight >20 kDa with a pI between 3 and 5 or 8 and 10 with an amount <50 ng in a series of experiments conducted <1 year ago. Such searches can quickly target potentially interesting molecules for further analysis.

With the increasing ease of transferring data through the Internet as well as local- and wide-area networks, it has become practical to quickly find and examine data from distant locations. Of course, great care must be taken to ensure that similar experimental conditions are employed, as otherwise the results will be difficult to compare. In this manner it is sometimes possible to dramatically increase the sample size and statistical accuracy as well as the probability of detecting rare events. In addition, if one data set is more completely characterized, this extra information can be extracted and applied to the other data set. For example if there is a band in common in two databases and there is sequence information for it in one database, that sequence information can be added to the other database. Currently most public electrophoresis database sites are two-dimensional protein databases. A list with links to many of these Internet database sites can be found at *http://www-lmmb.ncifcrf.gov/EP/table2Ddatabases.html*.

With biological questions becoming more complicated and the answers to the questions often requiring information from a variety of sources, it is becoming increasingly important to be able to move easily between information sources. A relational type of database can help achieve this. Unlike a conventional database with a fixed arrangement of data, relational databases have links between related files that allow for easy movement from one file to another. Another approach to interconnecting electrophoresis data with data from other sources is to generate a series of hypertext links between data sets, similar to what occurs on the Internet. Selecting a specific link moves the search to the related network site and the related information. Regardless of the method, the end goal is similar. An example of what is possible: a researcher selects a protein spot on a two-dimensional gel image, which triggers accessing of related information on this protein. The protein sequence is accessed from mass spectroscopy analysis of the spot on a separate gel. The sequence of the gene and the cDNA that generated the protein is retrieved. The expression pattern of the gene in various tissues and conditions, as well as information on similar genes in other organisms, is incorporated. Citations and annotations to this are retrieved as well. All of this information is compiled automatically into an interactive report about the protein. From this report, the researcher can formulate a more refined hypothesis and plan the most appropriate experiments to test it.

LITERATURE CITED

Appel, R.D., Hochstrasser, D.F., Funk, M., Vargas, J.R., Muller, A.F., and Scherrer, J.-R. 1991. The MELANIE project: From a biopsy to automatic protein map interpretation by computer. *Electrophoresis* 12:722-735.

Garrels, J.I. 1989. The QUEST system for quantitative analysis of two-dimensional gels. *J. Biol. Chem.* 264:5269-5282.

Glasbey, C.A. and Horgan, G.W. 1994. Image Analysis for the Biological Sciences. John Wiley & Sons, Chichester, England.

Hamming, R.W. 1973. Numerical methods for scientists and engineers, 2nd ed. Dover Publications, New York.

Huffman, D.A. 1952. A method for the construction of minimum-redundancy codes. *Proc. Inst. Elect. Radio Eng.* 40:9-12.

BIU WIEN CENTRAL UNIVERSITY LIBRARY

Monardo, P.J., Boutell, T., Garrels, J.I., and Latter, G.I. 1994. A distributed system for two-dimensional gel analysis. *Comput. Appl. Biosci.* 10:137-143.

Patton, W.F. 1995. Biologist's perspective on analytical imaging systems as applied to protein gel electrophoresis. *J. Chromatogr. A.* 698:55-87.

Plikaytis, B.D., Carlone, G.M., Edmonds, P., and Mayer, L.W. 1986. Robust estimation of standard curves for protein molecular weight and linear-duplex DNA base-pair number after gel electrophoresis. *Anal. Biochem.* 152:346-364.

Russ, J.C. 1995. The Image Processing Handbook. CRC Press, Boca Raton, Fla.

Smith, J.M. and Thomas, D.J. 1990. Quantitative analysis of one-dimensional gel electrophoresis profiles. *Comput. Appl. Biosci.* 6:93-99.

Sutherland, J.C., Lin, B., Monteleone, D.C., Mugavero, J., Sutherland, B.M., and Trunk, J. 1987. Electronic imaging system for direct and rapid quantitation of fluorescence from electrophoretic gels: Application to ethidium bromide–stained DNA. *Anal. Biochem.* 163:446-457.

Welch, T.A. 1984. A technique for high performance data compression. *IEEE Computer.* 17:21-32.

Sutherland, J.C. 1993. Electronic imaging of electrophoretic gels and blots. *In* Advances in Electrophoresis, Vol. 6. (A. Chrambach, M.J. Dunn, and B.J. Radola, eds.) pp. 1-41. VCH Verlagsgesellschaft mbH, Weinheim, Germany.

Provides an overview of image capture with particular emphasis on types of capture equipment.

KEY REFERENCES

Glasbey and Horgan, 1994. See above.

Describes general image-processing techniques as they are applied to biological images.

Russ, 1995. See above.

A general reference book on digital image capture and analysis.

INTERNET RESOURCES

rsb.info.nih.gov/nih-image

NIH Image is free software that provides basic image analysis tools for the Macintosh.

www.inforamp.net/~poynton/Poynton-color.html

Contains an excellent description of gamma correction in the Gamma FAQ.

www-immb.ncifcrf.gov/EP/table2Ddatabases.html

A list of links to many two-dimensional databases that are available via the Internet.

Contributed by Scott Medberry
Amersham Pharmacia Biotech
San Francisco, California

Sean Gallagher
Motorola Corporation
Tempe, Arizona

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Mass Changes Due to Some Post-Translational Modifications of Peptides and Proteins[a]

Modification	Average Mass Change
Pyroglutamic acid formed from Gln	−17.0306
Disulfide bond formation	−2.0159
C-terminal amide formed from Gly	−0.9847
Deamidation of Asn and Gln	−0.9847
Methylation	14.0269
Hydroxylation	15.9994
Oxidation of Met	15.9994
Proteolysis of a single peptide bond	18.0153
Formylation	28.0104
Acetylation	42.0373
Carboxylation of Asp and Glu	44.0098
Phosphorylation	79.9799
Sulfation	80.0642
Cysteinylation	119.1442
Pentoses (Ara, Rib, Xyl)	132.1161
Deoxyhexoses (Fuc, Rha)	146.1430
Hexosamines (GalN, GlcN)	161.1577
Hexoses (Fru, Gal, Glc, Man)	162.1424
Lipoic acid (amide bond to lysine)	188.3147
N-acetylhexosamines (GalNAc, GlcNAc)	203.1950
Farnesylation	204.3556
Myristoylation	210.3598
Biotinylation (amide bond to lysine)	226.2994
Pyridoxal phosphate (Schiff Base formed to lysine)	231.1449
Palmitoylation	238.4136
Stearoylation	266.4674
Geranylgeranlylation	272.4741
N-acetylneuraminic acid (Sialic acid, NeuAc, NANA, SA)	291.2579
Glutathionylation	305.3117
N-glycolylneuraminic acid (NeuGe)	307.2573
5′-Adenosylation	329.2091
4′-Phosphopantetheine	339.3294
ADP-ribosylation (from NAD)	541.3052

[a]To obtain the molecular mass of a modified peptide or protein, the appropriate mass changes should be algebraically added directly to the molecular mass calculated for the unmodified molecule.

HOW TO USE POST BINDERS

To open: pull tab at bottom of metal locking bar toward you until front cover releases. Pull out post bar to add or remove sheets.

To close: reinsert post bar, engage front cover, and push locking tab in completely.